

```

XX      SQ      Sequence      32 AA;
XX      Query Match      59.3%; Score 64; DB 21; Length 32;
XX      Best Local Similarity 72.2%; Pred. No. 0.0058;
XX      Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY      1 KMKLFKKIGIGKFLHSNK 18
XX      1 KMKLFKKIGIGKFLHSNK 18
Db      1 KMKLFKKIGIGKFLHSNK 18

RESULT 4
AAW97408
ID      AAW97408 standard; Protein; 39 AA.
XX
XX      AAW97408;
XX
XX      17-MAY-1999 (first entry)
XX
XX      Lytic peptide magainin 2+S.
XX
XX      magainin; transgenic; polinsetta; in vitro regeneration;
XX      resistance; ethylene; flower colour.
XX
XX      Synthetic.
XX
XX      WO9906566-A1.
XX
XX      11-FEB-1999.
XX
XX      31-JUL-1998; 98WO-US15917.
XX
XX      31-JUL-1997; 97US-0903944.
XX
XX      (SANF-) SANFORD SCI INC.
XX
XX      Blowers A, Chou T, Eisenreich R, Sanford J, Smith F;
XX      Van Eck J;
XX
XX      WPI; 1999-153796/13.
XX      N-PADB; AAX16060.
XX
XX      In vitro regeneration of Polinsetta - using embryo-induction medium
XX      containing casein hydrolase, used to produce transgenic plants
XX      with, e.g. increased disease resistance, controllable flower colour,
XX      altered habit and flowering time
XX
XX      Example 2; Page 51; 112pp; English.
XX
XX      The present sequence represents an lytic peptide which is used in the
XX      course of the invention to produce transgenic polinsetta plants with
XX      increased resistance to pathogens. The specification describes methods
XX      for the in vitro regeneration of Polinsetta. Transgenic Polinsetta can
XX      now be produced with e.g. increased resistance to pathogens or insects;
XX      controllable flower colour; modified habit (more compact form, earlier
XX      flowering) and/or increased resistance to ethylene.
XX
XX      SQ      Sequence      39 AA;
XX
XX      Query Match      59.3%; Score 64; DB 20; Length 39;
XX      Best Local Similarity 92.3%; Pred. No. 0.0071;
XX      Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY      8 IGIGKFLHSKKF 20
XX      :|||||
XX      16 MGIGKFLHSKKF 28
Db      16 MGIGKFLHSKKF 28

RESULT 5
AAW99125
ID      AAW99125 standard; peptide; 39 AA.
XX

```

```

AC      AAW99125;
XX
XX      17-MAY-1999 (first entry)
XX
XX      Extracellular localised antimicrobial magainin 2 peptide.
XX
XX      Antimicrobial peptide; resistance; plant pathogen; transgenic plant;
XX      powdery mildew.
XX
XX      Synthetic.
XX
XX      WO9906564-A1.
XX
XX      11-FEB-1999.
XX
XX      31-JUL-1998; 98WO-US16032.
XX
XX      31-JUL-1997; 97US-0054315.
XX
XX      (SANF-) SANFORD SCI INC.
XX
XX      Blowers AD, Sanford J, Smith F, Van Eck J;
XX      WPI; 1999-153794/13.
XX      N-PADB; AAX19274.
XX
XX      New antimicrobial peptides - comprising magainin and PGL classes of
XX      peptides having an N-terminal methionine, used particularly for
XX      providing resistance in plants to pathogens
XX
XX      Claim 5; Page 40; 52pp; English.
XX
XX      The present sequence represents an antimicrobial peptide. The present
XX      invention also describes a method of conferring microbial resistance on
XX      a plant cell comprising transforming the cell to express two peptides
XX      selected from the magainin and PGL classes, where the peptides are
XX      separately compartmentalised. The antimicrobial peptide is effective
XX      against phytopathogenic microbes including bacteria, fungi, and
XX      phytoplasma. It can also be used to target other plant pests, such as
XX      nematodes and viruses. The nucleic acid sequence encoding the peptide
XX      can be used to transform plants to provide antimicrobial resistance. The
XX      method is especially useful in preventing powdery mildew. The method
XX      allows for levels of peptide expression to provide resistance to
XX      phytopathogenic microbes without otherwise affecting the plant cells.
XX
XX      SQ      Sequence      39 AA;
XX
XX      Query Match      59.3%; Score 64; DB 20; Length 39;
XX      Best Local Similarity 92.3%; Pred. No. 0.0071;
XX      Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY      8 IGIGKFLHSKKF 20
XX      :|||||
XX      16 MGIGKFLHSKKF 28
Db      16 MGIGKFLHSKKF 28

RESULT 6
AAW67472
ID      AAW67472 standard; peptide; 19 AA.
XX
XX      AAW67472;
XX
XX      12-MAY-2000 (first entry)
XX
XX      Magainin 2 partial peptide fragment.
XX
XX      Antimicrobial; styelin; microbial degradation; bactericide; viricide;
XX      fungicide; plant protection; styelin A; magainin 2.
XX
XX      Unidentified.
XX
XX      CA2228730-A1.
XX

```

PD 24-JUN-1999.
 XX
 PF 14-APR-1998; 98CA-2228730.
 XX
 PR 24-DEC-1997; 97US-0068802.
 PR 20-JAN-1998; 98US-0072885.
 PR 18-FEB-1998; 98US-0075026.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Zhao C, Lee I, Lehrer RI;
 DR WPI; 2000-161486/15.
 XX
 PT Antimicrobial steylin peptides isolated from *Styela clava* useful for
 PT preserving materials vulnerable to microbial degradation and for
 PT protecting plants and animals against pathogenic microbes -
 XX
 PS Example 2; Page 22; 35pp; English.
 XX
 CC The invention relates to antimicrobial peptides (steylins) (and
 CC acetylated forms, salts, amides and esters of steylins), isolated from
 CC *Styela clava* and characterized by specific patterns of basic and
 CC hydrophobic amino acid side chains which display a broad spectrum of
 CC antimicrobial activities. The peptides are of a specified formula and
 CC display a wide range of antimicrobial activities and are therefore useful
 CC for preserving materials susceptible to microbial degradation, for
 CC protecting plants against bacterial infection and in the therapeutic and
 CC prophylactic protection of animals against bacteria, fungi and viruses.
 CC The compounds may also be used as standards in antimicrobial assays and
 CC as affinity ligands for absorption of counterpart structures in microbes,
 CC including viruses. The present sequence represents a partial fragment of
 CC magainin 2, used in comparison studies.
 XX
 SQ Sequence 19 AA;
 Query Match 58.3%; Score 63; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 0.0057;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GIGKFLHSARKF 20
 |||||
 Db 1 GIGKFLHSARKF 12
 RESULT 7
 AAP71208
 ID AAP71208 standard; Protein; 22 AA.
 XX
 AC AAP71208;
 XX
 DT 17-DEC-2001 (updated)
 DT 01-MAY-1991 (first entry)
 XX
 DE Magainin III polypeptide.
 XX
 KM Magainin; antimicrobial; antibiotic.
 XX
 OS synthetic.
 XX
 PN USN7021493-N.
 PD 08-SEP-1987.
 XX
 PF 04-MAR-1987; 87US-0021493.
 PR 04-MAR-1987; 87US-0021493.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN.
 XX
 PI Zaslloff MA;
 DR WPI; 1987-334845/47.

XX
 PT New antimicrobial polypeptide cpds. - are active against gram
 PT positive and negative bacteria, fungi and protozoa
 XX
 PS Disclosure; page 2; 11pp; English.
 XX
 CC This magainin III polypeptide has a mol. wt. of 25kD or less, is water
 CC sol. and non-cytolytic to animal cells, incl. RBCs, and is amphiphilic.
 CC It is an active antimicrobial agent being effective against both
 CC gram positive and -negative bacteria. It is also effective against
 CC fungi and protozoa. It is therefore useful as an antibiotic.
 CC See also AAP71206-07.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government-owned NPL applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dwpl/updates/ntls-us.html.)
 XX
 SQ Sequence 22 AA;
 Query Match 58.3%; Score 63; DB 8; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.0057;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GIGKFLHSARKF 20
 |||||
 Db 1 GIGKFLHSARKF 12
 RESULT 8
 AAP82736
 ID AAP82736 standard; Protein; 22 AA.
 XX
 AC AAP82736;
 XX
 DT 17-DEC-2001 (updated)
 DT 10-DEC-1990 (first entry)
 XX
 DE Magainin polypeptide (III) as bioactive cpds.
 XX
 KM Magainin polypeptide; cytotoxic agent; spermicidal agent;
 KM antimicrobial agent; therapeutic; antitumour activity; contraceptive;
 KM leukaemia; AIDS; HTLV III.
 XX
 OS synthetic.
 XX
 PN USN7076734-N.
 PD 23-FEB-1988.
 XX
 PF 23-JUL-1987; 87US-0076734.
 PR 23-JUL-1987; 87US-0076734.
 PR 04-MAR-1987; 87US-0021493.
 PR 05-AUG-1987; 87US-0081793.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN.
 XX
 PI Zaslloff MA;
 DR WPI; 1988-105363/15.
 XX
 PT New Magainin polypeptide(s) - active against Gram-positive and
 PT -negative bacteria, fungi, viruses and protozoa and as cytotoxic
 PT and spermicidal agents
 XX
 PS Disclosure; ; pp; English.
 XX
 CC To GI is attached NH2 and to N22 is attached OH.
 CC Magainins are a class of substantially pure, homogeneous
 CC polypeptides composed exclusively of about 25 amino acids, having a
 CC mol. wt. of about 2500 or less, being water soluble at a concentration
 CC of greater than 5 mg/ml at neutral pH or in an aq. solution of
 CC physiologic ionic strength. They are amphiphilic surface-seeking and have

CC a broad spectrum of properties at physiologic ionic strength and pH.
CC The polypeptide is a microbial agent capable of inhibiting the
CC growth or proliferation of gram-positive and gram-negative bacteria,
CC fungi, virus and protozoan species. They are also useful as
CC therapeutic cytotoxic agents. This includes antitumour activity.
CC They are useful as spermicides and as a contraceptive agent.
CC Magalain is active against HIV in vitro and has the potential
CC to reduce the infectivity of the AIDS virus indicating its use as a
CC sterilant against the virus or as a chemotherapeutic agent for treating
CC AIDS.
CC See also AAp82734-37.
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NIRS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpl/updates/ntis-us.html.)
XX Sequence 22 AA:

Query Match	58.3%	Score 63	DB 9	Length 22
Best Local Similarity	100.0%	Pred. No.	0.0057	
Matches 12	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```

OY      9 GIGKFLHSAKKF 20
         |||||
Db      1 GIGKFLHSAKKF 12

```

RESULT 9
AAR05732
ID AAR05732 standard; protein; 22 AA.
vv

DT 05-NOV-1990 (first entry)

Deletion analogue of Magalnin II.

KW Magalnin II; antihelminthic; spermicide; cancer; tumor;
KW fungicide; preservative.

Synthetic.

PN WO9006129-A.

PD 14-JUN-1990.

PF 05-DEC-1989; 89WO-0005499.

PR 09-DEC-1988; 88US-0281981.

PA (SCRI-) SCRIPPS CLINIC AND RE.

PI Cuervo JH, Houghten RA;
yy

DR WPI; 1990-209623/27.

PT Deletion analogues of Magainin I and II peptide(s) -
PT having high activity against microorganisms with low haemolytic
PT action against human red blood cells.

PS Claim 24; Page 24; 28pp; English.

CC analogues have the same activity as Megahalin I and II coupled with
CC insignificant haemolytic action against human red blood cells. Useful
CC as antibiotics, fungicides, spemticides, in destroying tumours and
CC as preservatives and in sterilisation of food etc.

50 **Sequence** **22** **AA;**

Query Match	58.3%	Score 63	DB 11	Length 22
Similarity	100.0%	Pred. NO.	0.0057	
Best Local				
Matches 12	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY      9 GIGKFLHSAKKF 20
          |||||
Db      1 GIGKFLHSAKKF 12
```

RESULT 10
AAR05733
ID AAR05733 standard; protein; 22 AA
vv

DT 05-NOV-1990 (first entry)
XX

Deletion analogue of Magalnin II.

KW Magainin II; antibiotic; spermicide; cancer; tumor;
KW fungicide; preservant.

OS Synthetic.

PN WO9006129-A.

PD 14-JUN-1990.

PF 05-DEC-1989; 89WO-0005499.

PR 09-DEC-1988; 88US-0281981.

PA (SCRI-) SCRIPPS CLINIC AND RE.

PI Cuervo JH, Houghten RA;

DR WPI; 1990-209623/27.

PT Deletion analogues of Magainin I and II peptide(s) -
PT having high activity against microorganisms with low haemolytic
PT action against human red blood cells.

PS Claim 25; Page 24; 28pp; English

CC Analogous have the same activity as Megahnin I and II coupled with
CC insignificant haemolytic action against human red blood cells. Useful
CC as antibiotics, fungicides, spermicides, in destroying tumors and
CC as preservatives and in sterilisation of food etc.

Sequence 22 AA;

Query Match	58.3%	Score 63	DB 11	Length 22
Best Local Similarity	100.0%	Pred. No.	0.0057	
Matches 12; Conservative	0;	Mismatches	0;	Gaps 0;

```
QY      9 GIGKFLHSAKKF 20
          |||||
Db      1 GIGKFLHSAKKF 12
```

```

RESULT 11
AAR05734
ID AAR05734 standard; protein; 22 AA

```

DT 05-NOV-1990 (first entry)

DE Deletion analogue of Magalnin II.

KW Magalnin II; antibiologic; spermicide; cancer; tumor;
KW fungicide; preservant.

05 Synthetic.

PN W09006129-A

PD 14-JUN-1990.
 XX
 PF 05-DEC-1989; 89WO-0005499.
 XX
 PR 09-DEC-1988; 88US-0281981.
 XX
 PA (SCRI-) SCRIPPS CLINIC AND RE.
 XX
 PI Cuervo JH, Houghten RA;
 XX
 DR WPI; 1990-209623/27.
 XX
 XX Deletion analogues of Magainin I and II peptide(s) -
 PT having high activity against microorganisms with low haemolytic
 PT action against human red blood cells.
 XX
 PS Claim 26; Page 24; 28pp; English.
 XX
 CC Analogues have the same activity as Magainin I and II coupled with
 CC insignificant haemolytic action against human red blood cells. Useful
 CC as antibiotics, fungicides, spermicides, in destroying tumors and
 CC as preservatives and in sterilisation of food etc.
 XX
 SQ Sequence 22 AA;
 Query Match 58.3%; Score 63; DB 11; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.0057;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GIGKFLHSAKFF 20
 DB 1 GIGKFLHSAKFF 12
 RESULT 12
 AAR05731
 ID AAR05731 standard; protein; 22 AA.
 XX
 AC AAR05731;
 XX
 DT 05-NOV-1990 (first entry)
 XX
 DE Deletion analogue of Magainin II.
 XX
 KW Magainin II; antibiotic; spermicide; cancer; tumor;
 KW fungicide; preservative.
 XX
 OS Synthetic.
 XX
 PN WO9006129-A.
 XX
 PD 14-JUN-1990.
 XX
 PF 05-DEC-1989; 89WO-0005499.
 XX
 PR 09-DEC-1988; 88US-0281981.
 XX
 PA (SCRI-) SCRIPPS CLINIC AND RE.
 XX
 PI Cuervo JH, Houghten RA;
 XX
 DR WPI; 1990-209623/27.
 XX
 XX Deletion analogues of Magainin I and II peptide(s) -
 PT having high activity against microorganisms with low haemolytic
 PT action against human red blood cells.
 XX
 PS Claim 23; Page 24; 28pp; English.
 XX
 CC Analogues have the same activity as Magainin I and II coupled with
 CC insignificant haemolytic action against human red blood cells. Useful
 CC as antibiotics, fungicides, spermicides, in destroying tumors and
 CC as preservatives and in sterilisation of food etc.

XX
 SQ Sequence 22 AA;
 Query Match 58.3%; Score 63; DB 11; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.0057;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GIGKFLHSAKFF 20
 DB 1 GIGKFLHSAKFF 12
 RESULT 13
 AAR10658
 ID AAR10658 standard; protein; 22 AA.
 XX
 AC AAR10658;
 XX
 DT 18-APR-1991 (first entry)
 XX
 DE Magainin II analogue (XII).
 XX
 KW Magainin I; magainin II; substitution analogue; spermicide;
 KW antibiotic; antiviral; antitumor.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 18..18
 FT /label- E-F-moc-lysine
 XX
 PN WO9100869-A.
 XX
 PD 24-JAN-1991.
 XX
 PF 28-JUN-1990; 90WO-U003675.
 XX
 PR 07-JUL-1989; 89US-0376754.
 XX
 PA (SCRI-) SCRIPPS CLINIC & RE.
 XX
 PI Houghten RA, Cuervo JH;
 XX
 DR WPI; 1991-051314/07.
 XX
 XX Substitution analogues of magainin I and II - having increased
 PT biological activity, useful as spermicide, antibiotic, antiviral
 PT and antitumor agents
 XX
 PS Claim 26; Page 55; 58pp; English.
 XX
 CC The substitution analogue has biological activity equal to or
 CC greater than the parent Magainin II peptide. The analogue is
 CC effective as antibiotic and can be used to inhibit, prevent or
 CC destroy the growth or proliferation of microbes such as Gram-
 CC positive and Gram-negative bacteria, fungi, viruses and protozoa.
 CC It can also be used as a spermicide to inhibit, prevent or
 CC destroy the motility of sperm and as anti-tumour agent to inhibit
 CC the growth of or destroy tumours.
 CC See also AAR10639-47, AAR10657-63 and AAR10999.
 XX
 SQ Sequence 22 AA;
 Query Match 58.3%; Score 63; DB 12; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.0057;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GIGKFLHSAKFF 20
 DB 1 GIGKFLHSAKFF 12
 RESULT 14


```

AARI0641
ID AARI0641 standard; Protein; 22 AA.
XX
XX AARI0641;
AC
XX
XX 18-APR-1991 (first entry)
DE Magalain II analogue (IV).
XX
XX Magalain I; magalain II; substitution analogue; spermicide;
KW antibiotic; antiviral; antitumour.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH MISC-difference 18..18
FT /label= G, A
XX
XX WO9100869-A.
XX
XX 24-JAN-1991.
XX
XX 28-JUN-1990; 90MO-U003675.
XX
XX 07-JUL-1989; 89US-0376754.
XX
XX (SCRI-) SCRIPPS CLINIC & RE.
XX
XX Houghten RA, Cuervo JH;
XX
XX WPI; 1991-051314/07.
XX
XX Substitution analogues of magalain I and II - having increased
PT biological activity, useful as spermicide, antibiotic, antiviral
PT and antitumour agents
XX
XX
XX Claim 26; Page 55; 58pp; English.
XX
XX The substitution analogue has biological activity equal to or
CC greater than the parent Magalain II peptide. The analogue is
CC effective as antibiotic and can be used to inhibit, prevent or
CC destroy the growth or proliferation of microbes such as Gram-
CC positive and Gram-negative bacteria, fungi, viruses and protozoa.
CC It can also be used as a spermicide to inhibit, prevent or
CC destroy the motility of sperm and as anti-tumour agent to inhibit
CC the growth of or destroy tumours.
CC See also AARI0639-47, AARI0657-63 and AARI0999.
XX
XX
SQ Sequence 22 AA;

Query Match 58.3%; Score 63; DB 12; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.0057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 GIGKFLHSARKF 20
DB 1 GIGKFLHSARKF 12

RESULT 15
AARI0642
ID AARI0642 standard; Protein; 22 AA.
XX
XX AARI0642;
AC
XX
XX 18-APR-1991 (first entry)
DE Magalain II analogue (V).
XX
XX Magalain I; magalain II; substitution analogue; spermicide;
KW antibiotic; antiviral; antitumour.
XX
XX Synthetic.
OS

```

```

XX
XX Key Location/Qualifiers
FH Modified-site 22..22
FT /note= "D-Serine"
XX
XX WO9100869-A.
XX
XX 24-JAN-1991.
XX
XX 28-JUN-1990; 90MO-U003675.
XX
XX 07-JUL-1989; 89US-0376754.
XX
XX (SCRI-) SCRIPPS CLINIC & RE.
XX
XX Houghten RA, Cuervo JH;
XX
XX WPI; 1991-051314/07.
XX
XX Substitution analogues of magalain I and II - having increased
PT biological activity, useful as spermicide, antibiotic, antiviral
PT and antitumour agents
XX
XX
XX Claim 26; Page 55; 58pp; English.
XX
XX The substitution analogue has biological activity equal to or
CC greater than the parent Magalain II peptide. The analogue is
CC effective as antibiotic and can be used to inhibit, prevent or
CC destroy the growth or proliferation of microbes such as Gram-
CC positive and Gram-negative bacteria, fungi, viruses and protozoa.
CC It can also be used as a spermicide to inhibit, prevent or
CC destroy the motility of sperm and as anti-tumour agent to inhibit
CC the growth of or destroy tumours.
CC See also AARI0639-47, AARI0657-63 and AARI0999.
XX
XX
SQ Sequence 22 AA;

Query Match 58.3%; Score 63; DB 12; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.0057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 GIGKFLHSARKF 20
DB 1 GIGKFLHSARKF 12

Search completed: June 6, 2003, 13:23:38
Job time : 31.5 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 13:22:11 ; Search time 12.5 Seconds

(without alignments)
153.815 Million cell updates/sec.

Title: US-10-081-418-1

Perfect score: 108

Sequence: 1 KMKLFKKIGKIFLHSAKAF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR_73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	58.3	303	2 A28620	magainin precursor
2	48	44.4	376	2 I40728	recombination prot
3	47.5	44.0	237	2 C75035	arginase related p
4	47.5	44.0	427	2 A84135	hypothetical prote
5	47.5	44.0	491	2 T50950	hypothetical prote
6	47.5	44.0	491	2 T50950	ATP-binding multid
7	46	42.6	412	2 S22474	phycoene synthase
8	46	42.6	412	2 A42102	phycoene synthase
9	46	42.6	666	2 G90566	hypothetical prote
10	45	41.7	64	1 CKMKA	cecropin A precurs
11	45	41.7	925	2 AE3628	nitrate reductase
12	44	40.7	205	2 C89957	hypothetical prote
13	44	40.7	353	2 S65297	probable membrane
14	44	40.7	512	2 H70662	probable p1cA prot
15	43	39.8	35	1 CKAOBP	cecropin B - chine
16	43	39.8	170	2 A75513	DNA topology modu
17	43	39.8	287	2 C71135	probable dihydrol
18	43	39.8	305	2 T27930	hypothetical prote
19	43	39.8	341	2 T16873	hypothetical prote
20	43	39.8	565	1 VHIYCB	nucleoprotein - in
21	43	39.8	589	2 T24378	hypothetical prote
22	43	39.8	620	2 F64408	coenzyme F420 hydr
23	43	39.8	662	2 UC7803	penicillin-binding
24	43	39.8	787	2 C75068	probable beta-gala
25	43	39.8	820	2 T26272	hypothetical prote
26	42.5	39.4	236	2 H82892	conserved hypothet
27	42	38.9	62	1 CKMKB	cecropin B precurs
28	42	38.9	63	1 CKMRA	cecropin A precurs
29	42	38.9	187	2 E84448	hypothetical prote

30	42	38.9	190	2 A64039	hypothetical prote
31	42	38.9	204	2 D69097	hypothetical prote
32	42	38.9	205	2 C95351	Vi-B5 type IV secr
33	42	38.9	325	2 A56406	casein kinase I (E
34	42	38.9	337	2 A57011	protein kinase I-1
35	42	38.9	337	2 S46254	hypothetical prote
36	42	38.9	380	2 T19096	probable phenylala
37	42	38.9	544	2 S75388	threonine ammonia-
38	42	38.9	600	2 T39516	hypothetical prote
39	42	38.9	5105	2 T32650	ubiquitinone/menaqu
40	41.5	38.4	248	2 F72262	A/G-specific adenil
41	41.5	38.4	339	2 B81258	serine/threonine/t
42	41.5	38.4	410	1 S28615	argininosuccinate
43	41.5	38.4	459	2 F84935	AAA family ATPase
44	41.5	38.4	809	2 T40537	hypothetical prote
45	41	38.0	106	2 T05199	

ALIGNMENTS

RESULT 1

A28620

Magainin precursor - African clawed frog

N:Contains: magainin 1; magainin 2

C:Species: Xenopus laevis (African clawed frog)

C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 18-Aug-2000

C:Accession: A28620; A29771

R:Entry, A.S.; Poulter, L.; Williams, D.H.; Nutkins, J.C.; Giovannini, M.G.; Moore, C

J. Biol. Chem. 263, 5745-5751, 1988

A>Title: The cDNA sequence coding for prepro-PGS (prepro-magainin) and aspects of th

A:Reference number: A28620; PMID:88186892; PMID:2833514

A:Accession: A28620

A:Molecule type: mRNA

A:Residues: 1-303 <TER>

A:Cross-References: GB:J03193; NID:9214654; PIDN:AAA49930.1; PID:9214655

R:Dasloff, M.

Proc. Natl. Acad. Sci. U.S.A. 84, 5449-5453, 1987

A>Title: Magainins, a class of antimicrobial peptides from Xenopus skin: Isolation, c

A:Reference number: A29771; PMID:87261003; PMID:3299384

A:Accession: A29771

A:Molecule type: mRNA

A:Residues: 6-73, 'Q', '75-158, 297-303 <ZAS>

C:Superfamily: magainin precursor

Query Match

Best Local Similarity 58.3%; Score 63; DB 2; Length 303;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GIGKFLHSAKAF 20

DB 83 GIGKFLHSAKAF 94

RESULT 2

I40728

recombination protein recA - Corynebacterium glutamicum

N:Alternate names: recombinase A

C:Species: Corynebacterium glutamicum

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Apr-2001

C:Accession: I40728; S41939; I40727; S39258

R:Billman-Jacobe, H.

DNA Seq. 4, 403-404, 1994

A>Title: Nucleotide sequence of a recA gene from Corynebacterium glutamicum.

A:Reference number: I40728; PMID:95143581; PMID:7841463

A:Accession: I40728

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-376 <BIU>

A:Cross-References: EMBL:X77384; NID:g475063; PIDN:CAA54563.1; PID:g475064

A:Experimental source: strain AS019

R:Billman-Jacobe, H

submitted to the EMBL Data Library, January 1994

```
AB4155
hypothetical protein BH4041 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Cdate: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: AB4155
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:title: Complete genome sequence of the alkaliphillic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: AB4155
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-427 <STO>
A:Cross-references: GB:A0001520; GB:BA000004; NID:g10176401; PIDN:BAB07760.1; GSPPDB:c.c
A:Experimental source: strain C-125
A:Gene: BH4041

Query Match      44.0%; Score 47.5; DB 2; Length 427;
Best Local Similarity 45.8%; Pred.No.6.4;
Matches 11; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

QY    1 KWKLFRKIGIC-----KFHSAKK 19
Db     217 WKPFVKKRGVGIRATKLVERAKK 240
        ||| | | :|| 
        ||| | | :|| 

RESULT 5
T50950
hypothetical protein B24P7.50 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Cdate: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 08-Sep-2000
C:Accession: T50950
R:Schultze, U.; Alqn, V.; Hohelisel, J.; Brandt, P.; Farmann, B.; Holland, R.; Nyakatutu
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T50950
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-491 <SCH>
A:Cross-references: EMBL:AL389890; GSPPDB:GN00116; NCSP:B24P7.50
A:Experimental source: BMC clone B24P7; strain OR74A
C:Genetics:
A:Gene: NCSP:B24P7.50
A:Map position: 6
A:introns: 30/3; 235/2
C:Superfamily: Neurospora crassa hypothetical protein B24P7.50

Query Match      44.0%; Score 47.5; DB 2; Length 491;
Best Local Similarity 50.0%; Pred.No.7.4;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY    2 WKLFR-KIIGIKFLHSAKP 20
Db     269 WTLMKCPSGCHKYMNSAEKF 288
        |:| |   |:::||:|
        |:| |   |:::||:|

RESULT 6
T43022
ATP-binding multidrug cassette transport protein - Botryotinia fuckelliana
C:Species: Botryotinia fuckelliana
C:Cdate: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T43022
R:Del Sorbo, G.; Andrade, A.C.; Van Nistelrooy J, G.M.; Van Kan J, A.L.; Balzi, E.; D
Mol. Gen. Genet. 254, 417-426, 1997
A:title: Multidrug resistance in Aspergillus nidulans involves novel ATP-binding cass
A:Reference number: Z07910; MUID:9180695; PMID:9180695
A:Accession: T43022
A>Status: preliminary; translated from GB/EMBL/DDBU
A:Molecule type: DNA
A:Residues: 1-1562 <EL>
A:Cross-references: EMBL:Z68906; NID:e989397; PID:e219958; PIDN:CAA93142.1
A:Experimental source: strain SAS56; clone bc-atra
```

C:Genetics:

A:Introns: 188/3; 338/3; 655/3; 711/2; 814/1; 1356/3
 C:Superfamily: unassigned ATP-binding cassette proteins: ATP-binding cassette homology

Query Match

Best Local Similarity 43.5%; Score 47; DB 2; Length 1562;
 Matches 8; Conservative 7; Mismatches 3; Indels 8; Gaps 1;

OY 1 KWK-----LKKIGIGKFLHSAK 18
 DB 1497 KWRNFGIFGLYQDLGVLGHQOR 1522

phytoene synthase (EC 2.5.1.-) - tomato
 N:Alternate names: ripening protein 5
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 02-Mar-2001
 C:Accession: S22474; S06321; S35155; S35154; S21981; S24967; S24968
 R:Ray, J.; Moutereau, P.; Bird, C.; Bird, A.; Grierson, D.; Maunders, M.; Truesdale, M.; H
 Plant Mol. Biol. 19, 401-404, 1992
 A:Title: Cloning and characterization of a gene involved in phytoene synthesis from toma
 A:Reference number: S22474; MUID:9232971; PMID:1623189
 A:Accession: S22474

RESULT 7

A:Molecule type: DNA
 A:Residues: 1-412 <RAY>

A:Cross-references: EMBL:X60441; GB:S39572; NID:919248; PIDN:CAA42969.1; PID:919249
 A:Note: The authors did not translate the codon for residue 155
 R:Ray, J.; Bird, C.; Maunders, M.; Grierson, D.; Schuch, W.
 Nucleic Acids Res. 15, 10587, 1987

A:Title: Sequence of PTOM5, a ripening related cDNA from tomato.
 A:Reference number: S06321; MUID:88096591; PMID:1697097
 A:Accession: S06321

A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-406, 'N', 408-412 <RAY1>

A:Cross-references: GB:X00521; NID:919340; PIDN:CAA68575.1; PID:919341
 R:Frery, R.G.; Grierson, D.
 Plant Mol. Biol. 22, 589-602, 1993

A:Title: Identification and genetic analysis of normal and mutant phytoene synthase gene
 A:Reference number: S35154; MUID:93344508; PMID:8343597
 A:Accession: S35155

A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-387, 'NMLKDFSNFKSGNSMATTIVGLAPCE', 'TYIADRGPIGTF' <FRA1>

A:Cross-references: EMBL:X67144; NID:919346; PIDN:CAA47625.1; PID:919347
 A:Experimental source: mutant rx
 A:Accession: S35154

A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-107, 'VECLSDPDTERTGVSLPQVGVSGSGAT', 'RWPGGPGRRGAGCIGCIGPLIFC' <FRA2>

A:Cross-references: EMBL:X67143; NID:919396; PIDN:CAA47624.1; PID:919399
 A:Experimental source: mutant r

C:Genetics:

A:Gene: Psy1
 A:Introns: 138/1; 155/1; 212/3; 291/2; 355/3
 C:Superfamily: Mycobacterium marinum phytoene synthase

C:Keywords: chloroplast; transferase

Query Match

Best Local Similarity 42.6%; Score 46; DB 2; Length 412;
 Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

OY 1 KWKLFKKIGIGKFLHSAK 20
 DB 322 KWRIFMK---KQIHRAKRF 337

RESULT 8

phytoene synthase (EC 2.5.1.-) peripheral plastid membrane - tomato
 C:Species: Lycopersicon esculentum (tomato)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 02-Mar-2001
 C:Accession: A42102

R:Barley, G.E.; Viltanen, P.V.; Bacot, K.O.; Scolnik, P.A.
 J. Biol. Chem. 267, 5036-5039, 1992

A:Title: A tomato gene expressed during fruit ripening encodes an enzyme of the carot
 A:Reference number: A42102; MUID:92184738; PMID:1544888
 A:Accession: A42102

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-412 <BAR>

A:Cross-references: GB:M84744; NID:9170415; PIDN:AAA34153.1; PID:9170416
 A:Note: sequence extracted from NCBI backbone (NCBIN:87066, NCBI:87067)

C:Superfamily: Mycobacterium marinum phytoene synthase
 C:Keywords: chloroplast; transferase

Query Match 42.6%; Score 46; DB 2; Length 412;
 Best Local Similarity 45.0%; Pred. No. 11;
 Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

OY 1 KWKLFKKIGIGKFLHSAK 20
 DB 322 KWRIFMK---KQIHRAKRF 337

RESULT 9

hypothetical protein MYPU_4390 [imported] - Mycoplasma pulmonis (strain UAB CTRP)
 C:Species: Mycoplasma pulmonis
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
 C:Accession: G90566

R:Chambers, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer,
 Nucleic Acids Res. 29, 2145-2153, 2001
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
 A:Reference number: A98512; MUID:21267165; PMID:11353084
 A:Accession: G90566

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-666 <KUR>

A:Cross-references: GB:AL445566; PID:914089853; PIDN:CAAC13612.1; GSPDB:GN00153
 A:Experimental source: strain UAB CTRP
 C:Genetics:

A:Gene: MYPU_4390
 A:Genetic code: SGC3

Query Match 42.6%; Score 46; DB 2; Length 666;
 Best Local Similarity 47.4%; Pred. No. 18;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 KWKLFKKIGIGKFLHSAK 19
 DB 350 KKHIFKKIGLGYVIFEX 368

RESULT 10

CKMKA
 cecropin A precursor - cecropia moth
 N:Alternate names: cecropin C

C:Species: Hyalophora cecropia (cecropia moth)
 C:Date: 24-Sep-1981 #sequence_revision 04-Nov-1994 #text_change 18-Jun-1999
 C:Accession: A40420; S00208; A91121; A93260; A01769

R:Gundundson, G.H.; Lidholm, D.A.; Asling, B.; Gan, R.; Boman, H.G.
 J. Biol. Chem. 266, 11510-11517, 1991
 A:Title: The cecropin locus. Cloning and expression of a gene cluster encoding three
 A:Reference number: A40420; MUID:91268009; PMID:1711035
 A:Accession: A40420

A:Molecule type: DNA
 A:Residues: 1-64 <GUD>

A:Cross-references: GB:M63845; NID:9159196; PIDN:AAA29185.1; PID:9159197
 R:Lidholm, D.A.; Gundundson, G.H.; Xanthopoulos, K.G.; Boman, H.G.
 FEBS Lett. 226, 8-12, 1987

A:Title: Insect immunity: cDNA clones coding for the precursor forms of cecropins A a
 A:Reference number: S00208
 A:Accession: S00208

A:Molecule type: mRNA
 A:Residues: 1-64 <LID>
 A:Cross-references: EMBL:X06672; NID:99413; PIDN:CAA29871.1; PID:g671870
 R:Hallmark, D.; Engstrom, A.; Bennich, H.; Kapur, R.; Boman, H.G.
 Eur. J. Biochem. 127, 207-217, 1982
 A:Title: Insect immunity: isolation and structure of cecropin D and four minor antibacterial peptides
 A:Reference number: A91121; MUID:83053366; PMID:7140755
 A:Accession: A91121
 A:Molecule type: protein
 A:Residues: 27-63 <HUL>
 A:Note: cecropin C lacks the carboxyl-terminal blocking group
 R:Steiner, H.; Hallmark, D.; Engstrom, A.; Bennich, H.; Boman, H.G.
 Nature 292, 246-248, 1981
 A:Title: Sequence and specificity of two antibacterial proteins involved in insect immunity
 A:Reference number: A93260; MUID:81245158; PMID:7019715
 A:Accession: A93260
 A:Molecule type: protein
 A:Residues: 27-63 <STE>
 R:Andreu, D.; Merrifield, R.B.; Steiner, H.; Boman, H.G.
 Proc. Natl. Acad. Sci. U.S.A. 80, 6475-6479, 1983
 A:Title: Solid-phase synthesis of cecropin A and related peptides.
 A:Reference number: A54960; MUID:84042494; PMID:6579533
 A:Contents: annotation; chemical synthesis
 A:Note: protein carboxyl-terminal composition and amidation determined by mass spectrometry
 C:Comment: Cecropins are a family of basic peptides having lytic and antibacterial activity
 C:Genetics:
 A:Introns: 34/3
 C:Superfamily: cecropin
 C:Keywords: amidated carboxyl end; antibacterial; hemolymph
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-26/Domain: propeptide #status predicted <PRO>
 F:27-63/Product: cecropin A #status experimental <MAT>
 F:63/Modified site: amidated carboxyl end (Lys) (amide in mature form from following gly)

Query Match 41.7%; Score 45; DB 1; Length 64;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KMKLFKKI 8
 DB 27 KMKLFKKI 34

RESULT 11
 AE3628
 nitrate reductase (EC 1.7.99.4) [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Jun-2002
 C:Accession: AE3628
 R:DeLVecchio, V.G.; Kaput, R.J.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD352; PMID:11756688
 A:Accession: AE3628
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-925 <KUR>
 A:Cross-references: GB:AE008918; PIDN:AAL54192.1; PID:g17985160; GSPDB:GN00191
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME110950
 A:Map position: 11
 C:Keywords: oxidoreductase

Query Match 41.7%; Score 45; DB 2; Length 925;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 KMKLFKKIGKFLHSKRF 20
 DB 72 KMKLFKKIGKFLHSKRF 91

RESULT 12
 C89957
 hypothetical protein SA1548 [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: C89957
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cul, L.; O ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: C89957
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-205 <KUR>
 A:Cross-references: GB:BA000018; PID:g13701522; PIDN:BA842816.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: SA1548

Query Match 40.7%; Score 44; DB 2; Length 205;
 Best Local Similarity 55.0%; Pred. No. 12;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 KMKLFKKIGKFLHSKRF 20
 DB 63 KMKLFKKIGKFLHSKRF 82

RESULT 13
 S65297
 probable membrane protein YPL264c - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein P0373
 C:Species: Saccharomyces cerevisiae
 C>Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 19-Apr-2002
 C:Accession: S65297; S65318
 R:Duesterhoeft, A.; Floeth, M.; Filtz, M.; Hilbert, H.; Moestl, D.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S65292
 A:Accession: S65297
 A:Molecule type: DNA
 A:Residues: 1-353 <DUE>
 A:Cross-references: EMBL:Z73620; NID:g1370544; PID:e246970; PID:g1370545; MIPS:YPL264
 A:Experimental source: strain S288C (AB972)
 R:Delius, H.; Hebling, U.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64967
 A:Accession: S65318
 A:Molecule type: DNA
 A:Residues: 1-353 <DUE>
 A:Cross-references: EMBL:Z73620; NID:g1370544; PID:e246970; PID:g1370545; MIPS:YPL264
 A:Experimental source: strain S288C (AB972)
 C:Genetics:
 A:Cross-references: SGD:S0006185
 A:Map position: 16L
 C:Keywords: transmembrane protein
 F:16-33/Domain: transmembrane #status predicted <TM1>
 F:54-70/Domain: transmembrane #status predicted <TM2>
 F:96-112/Domain: transmembrane #status predicted <TM3>
 F:118-134/Domain: transmembrane #status predicted <TM4>
 F:143-159/Domain: transmembrane #status predicted <TM5>
 F:189-205/Domain: transmembrane #status predicted <TM6>
 F:224-240/Domain: transmembrane #status predicted <TM7>

Query Match 40.7%; Score 44; DB 2; Length 353;
 Best Local Similarity 46.7%; Pred. No. 20;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 KMKLFKKIGKFLH 15
 DB 253 QWGLFNLIGISGFH 267

RESULT 14

H70662
 Probable pica protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: H70662
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Feldylin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: H70662
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-512 <COL>
 A:Cross-references: GB:283860; GB:AL123456; NID:93261681; PIDN:CAB06148.1; PID:91781258
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: pica

Query Match 40.7%; Score 44; DB 2; Length 512;
 Best Local Similarity 33.3%; Pred. No. 30;
 Matches 5; Conservative 8; Mismatches 2; Indels 0; Gaps 0;
 QY 2 WKLEKKGIGKFLHS 16
 ||:::||::||
 Db 233 WKYKNGKGLGRFINT 247

RESULT 15

CKROBP

cecropsin B - Chinese oak silkworm
 C:Species: Anthraea pernyi (Chinese oak silkworm)
 C:Date: 17-Dec-1982 #sequence_revision 12-Apr-1996 #text_change 07-May-1999
 C:Accession: A54725; A01771
 R:Craig, A.G.; Engstrom, A.; Bennich, H.; Kamensky, I.
 Blomd. Environ. Mass Spectrom. 14, 669-673, 1987
 A:Title: Plasma desorption mass spectrometry coupled with conventional peptide sequencing
 A:Reference number: A54725; MUID:88108273; PMID:2962676
 A:Accession: A54725
 A:Molecule type: protein
 A:Residues: 1-35 <CRA>
 R:Ou, X.; Steiner, H.; Engstrom, A.; Bennich, H.; Boman, H.G.
 Eur. J. Biochem. 127, 219-224, 1982
 A:Title: Insect immunity: isolation and structure of cecropsins B and D from pupae of the
 A:Reference number: A91122; MUID:83053368; PMID:6754375
 A:Accession: A01771
 A:Molecule type: protein
 A:Residues: 1-35, 'S' <QUX>
 A>Note: the carboxyl end is blocked and was not sequenced directly
 C:Superfamily: cecropsin
 C:Keywords: amidated carboxyl end; antibacterial
 F:35/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 39.8%; Score 43; DB 1; Length 35;
 Best Local Similarity 87.5%; Pred. No. 2.7;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKLEKKT 8
 ||::||
 Db 1 WKIKFKKT 8

Search completed: June 6, 2003, 13:25:26
 Job time : 14.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 13:21:00 ; Search time 7 Seconds

(without alignments)
118.504 Million cell updates/sec

Title: US-10-081-418-1

Perfect score: 108

Sequence: 1 KMKLFKKIGIGKFLHSKKF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	58.3	303	1	MAGL_XENLA
2	46	44.4	376	1	RECA_CORGL
3	48	42.6	412	1	PSY1_LYCES
4	45	41.7	64	1	CECA_HYACE
5	44	40.7	520	1	PHLA_MYCTU
6	43	39.8	35	1	CECB_ANTPE
7	43	39.8	341	1	YKLI_GAEEL
8	43	39.8	363	1	MTD3_STYHU
9	43	39.8	565	1	VNIC_INCCA
10	43	39.8	620	1	Y870_MERTA
11	42.5	39.4	713	1	GALX_SCHPO
12	42	38.9	62	1	CECB_HYACE
13	42	38.9	63	1	CECB_HYACE
14	42	38.9	190	1	YC31_HAEIN
15	42	38.9	325	1	KCIA_BOVIN
16	42	38.9	325	1	KCIA_BOVIN
17	42	38.9	337	1	KCIA_RAT
18	42	38.9	337	1	KCIA_CHICK
19	42	38.9	337	1	KCIA_HUMAN
20	42	38.9	544	1	STPB_SULSO
21	41.5	38.4	410	1	NKX3_HUMAN
22	41.5	38.4	459	1	APRA_ARATH
23	41.5	38.4	459	1	ARLY_BUCAT
24	41	38.0	305	1	BLAI_HAETN
25	41	38.0	336	1	KCIA_HAETN
26	41	38.0	552	1	KCIA_BOVIN
27	41	38.0	624	1	STI_SULAC
28	41	38.0	645	1	NKX3_RAT
29	41	38.0	713	1	NKX3_MOUSE
30	40.5	37.5	690	1	CN8A_HUMAN
31	40	37.0	635	1	POLG_DENIC
32	40	37.0	635	1	PCD8_DROME
33	40	37.0	635	1	CECA_BOMO
					CE3D_HYPCU
					CECB_BOMO

ALIGNMENTS

34	40	37.0	208	1	MLEY_HUMAN	P14649 homo sapien
35	40	37.0	247	1	RECO_NEIGO	Q9ZHY2 neisseria 9
36	40	37.0	368	1	YIUS_YEAST	P40577 saccharomyc
37	40	37.0	662	1	PMY7_YEAST	O06644 saccharomyc
38	40	37.0	684	1	TC10_YEAST	P50273 saccharomyc
39	40	37.0	755	1	TR2M_AGR4	P04029 agrobacteri
40	40	37.0	755	1	TR2M_AGR4	P25017 agrobacteri
41	40	37.0	1066	1	PDPI_ARATH	Q91R25 arabidopsis
42	40	37.0	1663	1	HAPF_HUMAN	O60229 homo sapien
43	40	37.0	1919	1	HAPF_HUMAN	P97924 rattus norv
44	40	37.0	3038	1	TRIO_HUMAN	O75962 homo sapien
45	40	37.0	3390	1	POLG_DEN3	P27915 d genome po

RESULT 1
MAGL_XENLA
ID MAGL_XENLA STANDARD: PRT; 303 AA.

AC P11006;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Magalins precursor.
OS Xenopus laevis (African clawed frog).
OC Amphibia; Batrachia; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OX Xenopodinae; Xenopus.
RN [1]
RP MEDLINE-8186892; PubMed-2833514;
RX Terry A.S., Poultier L., Williams D.H., Nutkins J.C., Giovannini M.G.,
Moore K.S., Bevilacqua C.L., Brasseur M.M., Tomassini N., Turner K.,
Eck H., Zaslouff M.,
RT "The cDNA sequence coding for prepro-PGS (prepro-magalins) and
RT aspects of the processing of this prepro-polypeptide.";
RL J. Biol. Chem. 263:5745-5751(1988).
[2]
RP SEQUENCE OF 6-158 AND 297-303 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-87261003; PubMed-3299384;
RA Zaslouff M.;
RT "Magalins, a class of antimicrobial peptides from Xenopus skin:
RT isolation, characterization of two active forms, and partial cDNA
RT sequence of a precursor.";
RN Proc. Natl. Acad. Sci. U.S.A. 84:5449-5453(1987).
[3]
RP SEQUENCE OF MAGALINS I AND II.
RX TISSUE-Stomach;
RX MEDLINE-92011794; PubMed-1717472;
RX Moore K.S., Bevilacqua C.L., Brasseur M.M., Tomassini N., Turner K.,
Eck H., Zaslouff M.,
RT "Antimicrobial peptides in the stomach of Xenopus laevis.";
RL J. Biol. Chem. 266:19851-19857(1991).
[4]
RP STRUCTURE BY NMR OF MAGALIN II.
RX MEDLINE-94129391; PubMed-8298457;
RA Bechinger B., Zaslouff M., Opella S.J.;
RT "Structure and orientation of the antibiotic peptide magalin in
RT membranes by solid-state nuclear magnetic resonance spectroscopy.";
RL Protein Sci. 2:2077-2084(1993).
-1- FUNCTION: ANTIMICROBIAL PEPTIDES THAT INHIBIT THE GROWTH OF
CC OF PROTOZOA. MAGALINS ARE MEMBRANE LYtic AGENTS.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE STOMACH AND STORED IN A
CC NOVEL GRANULAR MULTINUCLEATED CELL IN THE GASTRIC MUCOSA. IT IS
CC STORED AS ACTIVE, PROCESSED PEPTIDES IN LARGE GRANULES WITHIN
CC THE GRANULAR GLAND SECRETIONS OF THE SKIN.
CC -1- SIMILARITY: BELONGS TO THE MAGALIN FAMILY OF ANTIMICROBIAL
CC PEPTIDES.
CC -1- DATABASE: NMBE-Protein Spotlight;
CC NOTE-Issue 7 of February 2001;
CC WWW="http://www.expasy.org/spotlight/articles/spl1007.html".

Query Match	Best Local Similarity	Score 63:	DB 1:	Length 303:
Matches 12:	Conservative	100.0%:	Pred. NO. 0.0068:	
QY 9 GIGKFLHSAKKF 20				
DB 83 GIGKFLHSAKKF 94				
<p>Query Match: 58.3%; Score 63: DB 1: Length 303:</p> <p>Best Local Similarity: 100.0%; Pred. NO. 0.0068;</p> <p>Matches 12: Conservative 0; Mismatches 0; Indels 0; Gaps 0</p>				
<p>RECA CORGL</p> <p>STANDARD: PRT: 376 AA.</p>				
AC P42442;				
DT 01-NOV-1995 (rel. 32, Created)				
DT 01-NOV-1995 (rel. 32, Last sequence update)				
DT 15-JUN-2002 (rel. 41, Last annotation update)				
DE RecA protein (Recombinase A).				
GN RECA OR GGL1955.				
OS Corynebacterium glutamicum (Brevibacterium flavum).				
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;				
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;				
OC Corynebacterium.				
OX NCBI_TaxId=1718;				
RN [1]				
RP SEQUENCE FROM N.A.				
RP STRAIN=ATCC 13059 / AS019;				
RA Kerins S.M., Fitzpatrick R., O'Donoghue M., Dunican L.K.;				

```

RN      Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
|
|  [12]
|  SEQUENCE FROM N.A. / AS019;
|  STRAIN-ATCC 13059 / AS019;
|  MEDLINE=95143581; PubMed=7841463;
|  Billam-Jacobe H.;
|  "Nucleotide sequence of a recA gene from Corynebacterium glutamicum.";
|  DNA Seq. 4:403-404(1994).
|
|  [13]
|  SEQUENCE FROM N.A.
|  STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;
|  Nakagawa S.;
|  "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
|  Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
|
|  [14]
|  SEQUENCE OF 118-200 FROM N.A.
|  STRAIN-ATCC 13059 / AS019;
|  MEDLINE=95134369; PubMed=7765733;
|  Fitzpatrick R., O'Donohue M., Joy J., Heery D.M., Dunican L.K.;
|  "Construction and characterization of recA mutant strains of
|  Corynebacterium glutamicum and Brevibacterium lactofermentum.";
|  Appl. Microbiol. Biotechnol. 42:575-580(1994).
|
|  -1- FUNCTION: CAN CATALYZE THE HYDROLYSIS OF ATP IN THE PRESENCE OF
|  SINGLE-STRANDED DNA. THE ATP-DEPENDENT UPTAKE OF SINGLE-STRANDED
|  DNA BY DUPLEX DNA, AND THE ATP-DEPENDENT HYBRIDIZATION OF
|  HOMOLOGOUS SINGLE-STRANDED DNAS. IT INTERACTS WITH LEXA CAUSING
|  ITS ACTIVATION AND LEADING TO ITS AUTOCATALYTIC CLEAVAGE.
|  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
|  -1- SIMILARITY: BELONGS TO THE RECA FAMILY.
|
|  -----
|  This SWISS-PROT entry is copyright. It is produced through a collaboration
|  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
|  the European Bioinformatics Institute. There are no restrictions on its
|  use by non-profit institutions as long as its content is in no way
|  modified and this statement is not removed. Usage by and for commercial
|  entities requires a license agreement (see http://www.isb-sib.ch/announce/
|  or send an email to license@isb-sib.ch).
|  -----
|
|  DR      EMBL; U14965; AAD12743.1; -.
|  DR      EMBL; X77384; CAA54563.1; -.
|  DR      EMBL; AP005280; BAB9348.1; -.
|  DR      EMBL; X75085; CAA52977.1; -.
|  DR      HSSP; P26345; IG19.
|  DR      InterPro: IPR003593; AAA_ATPase.
|  DR      InterPro: IPR001553; RecA.
|  DR      Pfam; PF00154; RECA.1.
|  DR      PRINTS; PR00142; RECA.
|  DR      ProDom; PD000229; RecA; 1.
|  DR      SMART; SM00382; AAA; 1.
|  DR      PROSITE; PS00321; RECA_1; 1.
|  DR      PROSITE; PSS0162; RECA_2; 1.
|  DR      PROSITE; PSS0163; RECA_3; 1.
|  KW      DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding.
|  FT      NP_BIND 78 85
|  FT      CONFLICT 2 85 ATP (By similarity).
|  FT      SEQUENCE 376 AA; 40227 MW; 5535924E1B549D5C CRC64;
|
|  Query Match 44.4%; Score 48; DB 1; Length 376;
|  Best Local Similarity 63.6%; Pred. No. 2.3;
|  Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
|
|  QY      3 KLEKKIGIGKF 13
|  1:||||:|:|:|
|  DB      334 KIEFKKIGVGY 344
|
|  RESULT 3
|  ID      PSY1_LYCES STANDARD; PRT; 412 AA.
|  AC      P08196;
|  DT      01-AUG-1988 (Rel. 08, Created)
|  DT      01-OCT-1994 (Rel. 30, Last sequence update)
|  DT      15-JUN-2002 (Rel. 41, Last annotation update)

```

DE phytoene synthase 1, chloroplast precursor (EC 2.5.1.-) (Fruit
ripening specific protein PTOM5).
GN PSY1 OR PTOM5.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
[1] SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=92184738; PubMed=1544888;
RA Bartley G.E., Vitanen P.V., Bacot K.O., Scolnik P.A.;
RT "A tomato gene expressed during fruit ripening encodes an enzyme of
the carotenoid biosynthesis pathway.";
RL J. Biol. Chem. 267:5036-5039(1992).
RN [2]
[2] SEQUENCE FROM N.A.
RX STRAIN=cv. Marmande;
MEDLINE=94071905; PubMed=8250898;
RA Roemer S., Hugueney P., Bouvier F., Camara B., Kuntz M.;
RT Expression of the genes encoding the early carotenoid biosynthetic
enzymes in Capsicum annum.";
RL Biochem. Biophys. Res. Commun. 196:1414-1421(1993).
RN [3]
[3] SEQUENCE FROM N.A.
RX STRAIN=cv. Alisa Craig;
MEDLINE=88096591; PubMed=3697097;
RA Ray J., Bird C.R., Maunders M., Grierson D., Schuch W.;
RT "Sequence of PTOM5, a ripening related cDNA from tomato.";
RL Nucleic Acids Res. 15:10587-10587(1987).
RN [4]
[4] SEQUENCE FROM N.A.
RX STRAIN=cv. Alisa Craig; TISSUE=leaf;
MEDLINE=92322971; PubMed=1623189;
RA Ray J., Moureau P., Bird C., Grierson D., Maunders M.,
RT Trueblade M., Bramley P., Schuch W.;
RT Cloning and characterization of a gene involved in phytoene
synthesis from tomato.";
RL Plant Mol. Biol. 19:401-404(1992).
CC -1- FUNCTION: CATALYZES THE REACTION FROM PREPHYTOENE DIPHOSPHATE
TO PHYTOENE.
CC -1- CATALYTIC ACTIVITY: 2 geranylgeranyl diphosphate = diphosphate +
prephytoene diphosphate.
CC -1- CATALYTIC ACTIVITY: Prephytoene diphosphate = diphosphate +
phytoene.
CC -1- PATHWAY: Carotenoid biosynthesis.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- DEVELOPMENTAL STAGE: IN SEEDLINGS AND IN LATE STAGES OF FRUIT
RIPENING.
CC -1- INDUCTION: BY FRUIT RIPENING.
CC -1- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M84744; AAA34153.1; -
DR EMBL, X00521; CA688575.1; -
DR EMBL, X60441; CAA42969.1; -
DR EMBL, A21360; CAA01348.1; -
DR PIR, S06321; S06321.
DR PIR, A42102; A42102.
DR InterPro: IPR002060; Squ/phyt_synthase.
DR Pfam: PF00494; SOS_PSY; 1.
DR PROSITE: PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR PROSITE: PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
KW Multifunctional enzyme; Carotenoid biosynthesis; Multigene family;
KW Isoprene biosynthesis; Transferase; Chloroplast; Transit peptide.

FT TRANSIT 1 129 CHLOROPLAST (POMENTAL).
FT CHAIN 130 412 PHYTOENE SYNTHASE 1.
FT CONFLICT 389 407 KLALPLPAKSKSVPPKKT ->
FT CONFLICT 389 405 OVDCTYCKICKISCASYKN (IN REF. 3).
FT FT KLALPLPAKSKSVPPKKT ->
SO SEQUENCE 412 AA; 46615 MW; DBED05945D6F4C9A CRC64;
Query Match 42.68; Score 46; DB 1; Length 412;
Best Local Similarity 45.08; Pred. No. 5.2;
Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;
Qy 1 KMKLFKKIGIGKPLHSKKR 20
Db 322 KMRIFWK-----KQIHRRKRF 337
RESULT 4
CECA_HYACE STANDARD; PRT; 64 AA.
ID CECA_HYACE
AC P01507;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cecropin A precursor (Cecropin C).
OS Hyalophora cecropia (Cecropia moth).
OC Insecta; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Dityrpsia; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Hyalophora. Bombycoidea; Saturniidae; Saturniinae; Attacini;
OX NCBI_TaxID=7123;
RN [1]
[1] SEQUENCE FROM N.A.
RX MEDLINE=91268009; PubMed=1711035;
RA Gudmundsson G.H., Lidholm D.-A., Aasling B., Gan R., Boman H.G.;
RT "The cecropin locus. Cloning and expression of a gene cluster
encoding three antibacterial peptides in Hyalophora cecropia.";
RL J. Biol. Chem. 266:11510-11517(1991).
RN [2]
[2] SEQUENCE FROM N.A.
RA Lidholm D.-A., Gudmundsson G.H., Xanthopoulos K.G., Boman H.G.;
RT "Insect immunity: cDNA clones coding for the precursor forms of
cecropins A and D, antibacterial proteins from Hyalophora cecropia.";
RL FEBS Lett. 226:8-12(1987).
RN [3]
[3] SEQUENCE OF 27-63.
RX MEDLINE=83053366; PubMed=7140755;
RA Hultmark D., Engstroem A., Bennich H., Kapur R., Boman H.G.;
RT "Insect immunity: Isolation and structure of cecropin D and four
minor antibacterial components from Cecropia pupae.";
RL Eur. J. Biochem. 127:207-217(1982).
RN [4]
[4] PRELIMINARY SEQUENCE OF 27-63.
RX MEDLINE=81245158; PubMed=7019715;
RA Steiner H., Hultmark D., Engstroem A., Bennich H., Boman H.G.;
RT "Sequence and specificity of two antibacterial proteins involved in
insect immunity.";
RL Nature 292:246-248(1981).
RN [5]
[5] STRUCTURE BY NMR.
RX MEDLINE=89088132; PubMed=3207693;
RA Holak T.A., Engstroem A., Kraulis P.J., Lindeberg G., Bennich H.,
RA Jones T.A., Gronenborn A.M., Clore G.M.;
RT "The solution conformation of the antibacterial peptide cecropin A: a
nuclear magnetic resonance and dynamical simulated annealing study.";
RL Biochemistry 27:7620-7629(1988).
CC -1- FUNCTION: CECROPINS HAVE LYTIC AND ANTIBACTERIAL ACTIVITY AGAINST
SEVERAL GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: A PROTEIN WITH THE SAME SEQUENCE AS CECROPIN A, BUT LACKING
THE CARBOXYL BLOCKING GROUP, WAS ISOLATED AND CALLED CECROPIN C BY
REF. 3 AUTHORS.
CC -1- SIMILARITY: BELONGS TO THE CECROPIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X06672; CAA29871.1; -
 DR EMBL; M63845; AAA29185.1; -
 DR PIR; A01769; CCKWA.
 DR PIR; A40420; A40420.
 DR PIR; S00208; S00208.
 DR InterPro: IPR003254; IIP_Cecropin.
 DR InterPro: IPR003254; IIP_Cecropin.
 DR Pfam: PF00272; Cecropin; 1.
 DR ProDom: PD00396; IIP_Cecropin; 1.
 DR PROSITE: PS00268; CECROPIN; 1.
 KW Insect immunity; Antibiotic; Hemolymph; Amulation; Multigene family;
 KW Signal.
 FT SIGNAL 1 22
 FT PROPEP 23 26 REMOVED BY A DIPEPTIDYLPEPTIDASE.
 FT CHAIN 27 63 CECROPIN A.
 FT MOD_RES 63 63 AMIDATION (G-64 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 64 AA: 6952 MW: 0B3AEAL5C32DB3A4 CRC64;

Query Match 41.7%; Score 45; DB 1; Length 64;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KMKLFKKI 8
 |||||
 Db 27 KMKLFKKI 34

RESULT 5
 PHLA_MYCTU STANDARD; PRT: 520 AA.
 ID 004001; 053408; 050772; 050560; 050771; 008223;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phospholipase C 1 precursor (EC 3.1.4.3) (MTP40 antigen).
 GN PLCA OR MPCA OR MTP40 OR MTP351C OR MTP2416 OR MTCY98.20C.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=96029721; PubMed=7591062;
 RA Leao S.C., Rocha C.L., Murrillo L.A., Parra C.A., Patirroyo M.E.;
 RT "A species-specific nucleotide sequence of Mycobacterium tuberculosis
 RT Escherichia coli.";
 RT Infect. Immun. 63:4301-4306(1995).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-H37RV;
 RX MEDLINE=96333368; PubMed=8757862;
 RA Johansen K.A., Gill R.E., Vasil M.L.;
 RT "Biochemical and molecular analysis of phospholipase C and
 RT phospholipase D activity in mycobacteria.";
 RT Infect. Immun. 64:3259-3266(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garrier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RT Nature 393:537-544(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishal W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 323-520 FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=91372944; PubMed=1909999;
 RA Parra C.A., Londono L.P., del Portillo P., Patirroyo M.E.;
 RT "Isolation, characterization, and molecular cloning of a specific
 RT Mycobacterium tuberculosis antigen gene: identification of a species-
 RT specific sequence.";
 RT Infect. Immun. 59:3411-3417(1991).
 RN [6]
 RP SEQUENCE OF 355-496 FROM N.A.
 RX MEDLINE=94129477; PubMed=8298518;
 RA Leao S.C.;
 RT "tuberculosis: new strategies for the development of diagnostic tests
 RT and vaccines.";
 RT Braz. J. Med. Biol. Res. 26:827-833(1993).
 RN [7]
 RP POLYMORPHISM.
 RX MEDLINE=97269224; PubMed=9114405;
 RA Vera-Cabrera L., Howard S.T., Laszlo A., Johnson W.M.;
 RT "Analysis of genetic polymorphism in the phospholipase region of
 RT Mycobacterium tuberculosis.";
 RT J. Clin. Microbiol. 35:1190-1195(1997).
 CC -1- FUNCTION: HYDROLASES SPHINGOMYELIN IN ADDITION TO
 CC PHOSPHATIDYLCHOLINE. PROBABLE VIRULENCE FACTOR IMPLICATED IN THE
 CC PATHOGENESIS OF M. TUBERCULOSIS AT THE LEVEL OF INTRACELLULAR
 CC SURVIVAL, BY THE ALTERATION OF CELL SIGNALING EVENTS OR BY DIRECT
 CC CYTOTOXICITY.
 CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
 CC diacylglycerol + choline phosphate.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED, AT THE EXTRACELLULAR
 CC SIDE.
 CC -1- MISCELLANEOUS: POLYMORPHISM WAS DISCOVERED IN THE PHOSPHOLIPASE
 CC PLCA/B/C REGION. SOME STRAINS APPEARED TO LACK BOTH PLCA AND PLCB
 CC GENES, WHILE OTHERS LACK ONLY PLCB.
 CC -1- MISCELLANEOUS: AN IS6110 INSERTION ELEMENT HAS BEEN FOUND IN
 CC STRAIN LCDC-194, INTERRUPTING THE PLCA GENE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PHOSPHOLIPASE C FAMILY.
 CC -1- CAUTION: REF. 5 AND REF. 6 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC SEVERAL FRAMESHIFTS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L11868; AAB59164.1; ALT_INT.
 DR EMBL; L11868; AAB59165.1; ALT_SEQ.
 DR EMBL; 049511; AAC18943.1; -
 DR EMBL; Z83860; CAB06148.1; ALT_INIT.
 DR EMBL; AE007081; AAK46709.1; -

DR EMBL: M57952; AAA63288.1; ALT_FRAME.
 DR EMBL: S69737; AAC60465.2; ALT_FRAME.
 DR TIGR: MT2416; -
 DR Tuberculin; Signal; Membrane; Antigen; Complete proteome.
 KW Hydrolyase; Signal; Membrane; Antigen; Complete proteome.
 FT SIGNAL 1
 FT CHAIN 38 520
 FT CONFLICT 327 327
 FT CONFLICT 369 370
 FT CONFLICT 406 407
 FT CONFLICT 516 520
 SO SEQUENCE 520 AA; 56138 MW; 5844EF0C064288A7 CRC64;

Query Match 40.7%; Score 44; DB 1; Length 520;
 Best Local Similarity 33.3%; Pred. No. 14;
 Matches 5; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

OY 2 WKLFKKGIGKFLHS 16
 DB 241 WKVYONKGLGRFINT 255

RESULT 6

CECB_ANTPE STANDARD; PRT; 35 AA.
 AC P01509;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cecropin B.
 OS Antheraea pernyi (Chinese oak silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pelegrypota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 OC Ditrysia; Bombycoidea; Saturniidae; Saturniinae; Saturniini;
 OC Antheraea.
 NCBI_Taxid-7119;
 RN [1]
 RP PARTIAL SEQUENCE.
 RX MEDLINE:83053368; PubMed-6754375;
 RA Qu X.-M., Steiner H., Engstrom A., Bensch H., Boman H.G.;
 RT "Insect immunity: Isolation and structure of cecropin B and D from
 RT pupae of the Chinese oak silk moth, *Antheraea pernyi* L.",
 RT Eur. J. Biochem. 127:219-224(1982).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE-88108273; PubMed-2962676;
 RA Craig A.G., Engstrom A., Bensch H., Kamensky I.;
 RT "Plasma desorption mass spectrometry coupled with conventional
 RT peptide sequencing techniques.",
 RT Biomed. Environ. Mass Spectrom. 14:669-673(1987).
 CC -1- FUNCTION: CECROPINS HAVE LYTIC AND ANTIBACTERIAL ACTIVITY AGAINST
 CC SEVERAL GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE CECROPIN FAMILY.
 CC PIR: A01771; CKA0BP.
 DR PIR: A54725; A54725.
 DR InterPro: IPR000875; Cecropin.
 DR InterPro: IPR003254; IIP-cecropin.
 DR Pfam: PF00272; Cecropin; 1.
 DR ProDom: PD003966; IIP-cecropin; 1.
 DR PROSITE: PS00268; CECROPIN; 1.
 KW Insect immunity; Antibiotic; Hemolymph; Amidation; Multigene family.
 FT MOD_RES 35 35
 SO SEQUENCE 35 AA; 3818 MW; 7AF6B0A10E16313 CRC64;

Query Match 39.8%; Score 43; DB 1; Length 35;
 Best Local Similarity 87.5%; Pred. No. 1.6;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WKLFKKT 8
 DB 1 WKLFKKT 8

RESULT 7

YKLL_CAEEL STANDARD; PRT; 341 AA.
 ID YKLL_CAEEL
 AC P42168;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative casein kinase I C03C10.1 in chromosome III (EC 2.7.1.-).
 GN C03C10.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_Taxid-6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Berts M.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CASEIN KINASE I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z35637; CAAB4685.1; -
 DR HSP: 006486; ICR1.
 DR WormPep: C03C10.1; CE00872.
 DR InterPro: IPR00719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 16 284
 FT NP_BIND 22 30
 FT BINDING 45 45
 FT ACT_SITE 135 135
 SO SEQUENCE 341 AA; 39037 MW; 17AD6E173551BABC CRC64;

Query Match 39.8%; Score 43; DB 1; Length 341;
 Best Local Similarity 61.5%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 WKLFKKGIGKFLHS 13
 DB 15 KYLLRKIGSGSF 27

RESULT 8

MTD3_STYHU STANDARD; PRT; 363 AA.
 ID MTD3_STYHU
 AC Q43138;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Probable mannitol dehydrogenase 3 (EC 1.1.1.255) (NAD-dependent
 DE mannitol dehydrogenase 3).
 GN CAD3.
 OS Stylosanthes humilis (Towns-ville stylo).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Stylosanthes.

```

RX RN [1] NCBI_TaxID=35628;
RN RN SEQUENCE FROM N.A.
RC STRAIN=cv. Paterson; TISSUE=stem;
RA Nourse J.P., Manners J.M., Curtis M.D., Abrahams S.L., Watson J.M.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: OXIDIZES MANNITOL TO MANNOSE. PROVIDES THE INITIAL STEP
CC BY WHICH TRANSLOCATED MANNITOL IS COMMITTED TO CENTRAL METABOLISM
CC AND, BY REGULATING MANNITOL POOL SIZE, IS IMPORTANT IN REGULATING
CC SALT TOLERANCE AT THE CELLULAR LEVEL (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: D-mannitol + NAD(+) = D-mannose + NADH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE A CINNAMYL-ALCOHOL
CC DEHYDROGENASE.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC EMBL; L36456; AAA74883.1; .
CC DR InterPro; IPR002328; ADH_zinc.
CC DR InterPro; IPR002085; Adh_zn_family.
CC DR Pfam; PF00107; adh_zinc; 1.
CC DR PROSITE; PS00059; ADH_ZINC; 1.
CC -----
CC KW Oxidoreductase; NAD; Zinc; Multigene family.
CC FT METAL 51 51 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 73 73 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 104 104 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 107 107 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 110 110 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 118 118 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 168 168 ZINC (CATALYTIC) (BY SIMILARITY).
CC SQ SEQUENCE 363 AA; 39551 MW; EBBB21724513D90 CRC64;
CC -----
CC CC Query Match 39.8%; Score 43; DB 1; Length 363;
CC Best Local Similarity 38.9%; Pred. No. 14;
CC Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
CC -----
CC Oy 1 KWLEFKKIGIGKPLHSK 18
CC | : | : | : | : | :
CC Db 88 KYKGDYGVGVFVESC 105
CC -----
CC RESULT 9
CC VNUC_INCCA STANDARD; PRT; 565 AA.
CC AC P08028;
CC DRT 01-AUG-1988 (Rel. 08, Created)
CC DRT 01-AUG-1988 (Rel. 08, Last sequence update)
CC DRT 01-MAY-1992 (Rel. 22, Last annotation update)
CC DE Nucleoprotein.
CC GN NP
CC OS Influenza C virus (strain C/California/78).
CC OS Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC OC Influenzavirus C.
CC OX NCBI_TaxID=11555;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=85170522; Pubmed=6532006;
CC RA Nakada S., Creager R.S., Krystal M., Palase P.;
CC RT *Complete nucleotide sequence of the Influenza C/California/78 virus
CC RT nucleoprotein gene.*;
CC RL Virus Res. 1:433-441(1984).
CC CC -1- FUNCTION: ENCAPSIDATES THE NEGATIVE STRAND VIRAL RNA.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M17700; AAA43798.1; -
CC
CC PIR: S07413; VHIVC8.
CC
KW Nucleoprotein
SQ SEQUENCE 565 AA; 63598 MW; 22C10CCB1A85AA84 CRC64;
OY
OY 1 KWLIFKIGIGIFLHSAK 18
OY 11 : 1 1111 : 11
Db 181 KWIEMKRGIGVLTADAK 198

RESULT 10
Y870.METUA
ID Y870.METUA STANDARD; PRT: 620 AA.
AC Q58280;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0870.
GN MJ0870.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_Taxid:2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake O., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Kerlavage A.R., Kinkness E.F., Weissman J.K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.M., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT Science 273:1058-1073(1996).
RL -1- SIMILARITY: TO COENZYME F420 HYDROGENASE BETA SUBUNIT.
CC -1- SIMILARITY: TO M.JANNASCHII MJ1349, MJ0725 AND MJ0551.
CC -1- SIMILARITY: THE C-TERMINAL DOMAIN IS A 4FE-4S/SIROHEME DOMAIN
CC FOUND IN NITRITE REDUCTASES (EC 1.6.6.4 AND EC 1.7.7.1) AND
CC SULFITE REDUCTASES (EC 1.8.1.2 AND EC 1.8.7.1).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U67531; AAB98876.1; -
CC
CC HSSP: Q45560; 1BWE.
CC
CC TIGR: MJ0870; -
CC
CC InterPro: IPR001450; 4FE4S_ferredoxin.
CC InterPro: IPR000660; Nir_Sir.
CC InterPro: IPR005117; Nir_sir_fer.
CC Pfam: PF00037; fer4; 3.
CC Pfam: PF01077; nir_sir; 1.
CC Pfam: PF03460; nir_sir_fer; 1.
CC PRINTS: PR00397; SIROHAEM.
CC PROSITE: PS00198; 4FE4S_FERREDOXIN; 2.

```

DR PROSITE; PS00365; NIR_SIR. 1.
 KW Hypothetical protein; Oxidoreductase; Heme; Iron-sulfur; 4Fe-4S;
 Complete proteome.
 FT METAL 428 428 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 434 434 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 468 468 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 472 472 IRON-SULFUR (4FE-4S) AND SIROHEME
 (BY SIMILARITY).
 SQ SEQUENCE 620 AA; 69793 MW; 9D71D2580D7D0BA8 CRC64;
 Query Match 39.8%; Score 43; DB 1; Length 620;
 Best Local Similarity 43.8%; Pred. No. 23;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 OY 4 LFKKIGIGKFLHSKR 19
 : : : : :
 Db 598 VMKRVGKFLKEVKE 613
 RESULT 11
 ID GALX_SCHPO STANDARD; PRT; 713 AA.
 AC 09HDJ3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GAL10 bifunctional protein [includes: UDP-glucose 4-epimerase
 (EC 5.1.3.2) (galactowaldenase); Aldose 1-epimerase (EC 5.1.3.3)
 (Mutarotase)].
 GN GAL10 OR SPBP282.12C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE-21648401; PubMed-11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moulé S., Mungall K., Murphy L., Niblett D., Odell C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Woodward J., Volckaert G., Aert R., Robben J., Grynopreuz B.,
 Weljens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fliczc C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 Goffeau A., Cadieu E., Dreano S., Gloux S., Leleaux V., Mottier S.,
 Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 Dague R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Shpkovski G.V., Ussery D., Schizosaccharomycetes pombe.;
 RA "The genome sequence of Schizosaccharomycetes pombe.";
 RT Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: Mutarotase converts alpha-aldose to the beta-anomer. It
 is active on D-glucose, L-arabinose, D-xylose, D-galactose, maltose
 and lactose (By similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-glucose -> UDP-galactose.
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose -> beta-D-glucose.
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose -> beta-D-glucose.
 CC -1- PATHWAY: Galactose metabolism; third step.
 CC -1- PATHWAY: Hexose metabolism.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE SUGAR
 CC EPIMERASE FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ALDOSE
 CC EPIMERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL; A512522; CAC21414.1; -.
 DR HSP; P09147; 1XEL.
 DR InterPro; IPR001823; Ald1-epimerase.
 DR InterPro; IPR001509; Epimerase, Dh.
 DR InterPro; IPR00205; NAD binding.
 DR Pfam; PF01263; Aldose-epim: 1.
 DR Pfam; PF01370; Epimerase: 1.
 DR TIGR; TIGR01179; gale; 1.
 DR PROSITE; PS00345; ALDOSE_1-EPIMERASE; 1.
 KW Multifunctional enzyme; Isomerase; NAD; Galactose metabolism.
 FT DOMAIN 1 350 GALACTOWALDENASE.
 FT DOMAIN 351 713 MUTAROTASE.
 FT NP_BIND 7 38 NAD (POTENTIAL).
 FT ACT_SITE 532 532 MUTAROTASE (POTENTIAL).
 SQ SEQUENCE 713 AA; 80666 MW; 51C8D9A0843A8556 CRC64;
 Query Match 39.4%; Score 42.5; DB 1; Length 713;
 Best Local Similarity 47.4%; Pred. No. 32;
 Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
 OY 1 KWKLKKGIGKFG---LHS 16
 : : : : :
 Db 354 KWKLKKGIGKFG---LHS 372
 RESULT 12
 ID CECB_HYACE STANDARD; PRT; 62 AA.
 AC P01508;
 DT 21-AUG-1986 (Rel. 01, Created)
 DT 15-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cecropin B precursor (immune protein p9).
 GN Hyalophora cecropia (Cecropia moth).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 OC Diptera; Bombycoidea; Saturniidae; Saturniinae; Attacini;
 OC Hyalophora.
 OX NCBI_TaxID=7123;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86165708; PubMed-3127206;
 RA Xanthopoulos G., Lee J.-Y., Gan R., Kockum K., Faye I., Boman G.;
 RT "The structure of the gene for cecropin B, an antibacterial immune
 RT protein from Hyalophora cecropia.";
 RL Eur. J. Biochem. 172:371-376(1988).
 CC [2]
 CC SEQUENCE FROM N.A.
 RX MEDLINE-85190473; PubMed-3857578;
 RA van Hofsten P., Faye I., Kockum K., Lee J.-Y., Xanthopoulos K.G.,
 Boman I.A., Boman H.G., Engstroem A., Andreu D., Merrifield R.B.;
 RT "Molecular cloning, cDNA sequencing, and chemical synthesis of
 RT cecropin B from Hyalophora cecropia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2240-2243(1985).
 CC [3]
 CC SEQUENCE FROM N.A.
 RX MEDLINE-86005745; PubMed-3840100;
 RA Boman H.G., Faye I., von Hofsten P., Kockum K., Lee J.-Y.,
 Xanthopoulos K.G., Bennich H., Engstroem A., Merrifield R.B.,
 Andreu D.;

RT "On the primary structures of lysozyme, cecropins and attacins from
 RT Hyalophora cecropia.";
 RL Dev. Comp. Immunol. 9:551-558(1985).
 RN [41]
 RP SEQUENCE OF 27-61.
 RA MEDLINE-61245158; PubMed-7019715;
 RX Seitelner H., Hultmark D., Engstrom A., Bennich H., Boman H.G.;
 RT "Sequence and specificity of two antibacterial proteins involved in
 insect immunity.";
 RT Nature 292:246-248(1981).
 RL NATURE 292:246-248(1981).
 CC -1- FUNCTION: CECROPINS HAVE LYMIC AND ANTIBACTERIAL ACTIVITY AGAINST
 CC SEVERAL GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE CECROPIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: M10309; AAA29184.1; -;
 CC DR EMBL: X07404; CA30306.1; -;
 CC DR EMBL: M34924; AAA29187.1; -;
 CC DR PIR: A01770; CKMKB.
 CC DR PIR: S02606; S02606.
 CC DR InterPro: IPR000875; Cecropin.
 CC DR InterPro: IPR003254; IIP-cecropin.
 CC DR Pfam: PF00272; cecropin; 1.
 CC DR ProDom: PD003996; IIP-cecropin; 1.
 CC DR PROSITE: PS00268; CECROPIN; 1.
 CC DR Insect Immunity: Antibiotic; Hemolymph; Amidation; Multigene family;
 CC Signal.
 CC KW SIGNAL. 1 22
 CC FT PROPEP 23 26 REMOVED BY A DIPEPTIDYLPEPTIDASE.
 CC FT CHAIN 27 61 CECROPIN B.
 CC FT MOD_RES 61 61 AMIDATION (G-62 PROVIDE AMIDE GROUP).
 CC SQ SEQUENCE 62 AA; 6750 MW; 65E309CEB2983BD61 CRC64;
 CC -----
 CC Query Match 38.9%; Score 42; DB 1; Length 62;
 CC Best Local Similarity 87.5%; Pred. No. 4;
 CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC OY 1 KWKLFKKT 8
 CC DB 27 KWKVFKKT 34
 CC -----
 CC RESULT 13
 CC CECA_BOMMO STANDARD; PRT; 63 AA.
 CC ID CECA_BOMMO
 CC AC 027239;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Cecropin A precursor.
 CC GN CECA.
 CC OS Bombyx mori (Silk moth).
 CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 CC OC Diptera; Bombycoidea; Bombycidae; Bombyx.
 CC OX NCBI_TaxID=7091;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=C108; TISSUE=Larval fat body;
 CC RA MEDLINE-94369101; PubMed-7765280;
 CC RA Yamao Y., Matsumoto M., Inoue K., Kawabata T., Morishima I.;
 CC RT "Cloning of cDNAs for cecropins A and B, and expression of the genes
 CC in the silkworm, Bombyx mori.";
 CC RL Biosci. Biotechnol. Biochem. 58:1476-1478(1994).
 CC -1- FUNCTION: CECROPINS HAVE LYMIC AND ANTIBACTERIAL ACTIVITY AGAINST

CC SEVERAL GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN FAT BODY AND HEMOCYTES.
 CC IS ALSO EXPRESSED IN MALPIGHIAN TUBE AND TO A MUCH LESSER EXTENT
 CC IN MIDGUT. NOT PRESENT IN SILK GLAND.
 CC -1- SIMILARITY: BELONGS TO THE CECROPIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: D17394; BAA04217.1; -;
 CC DR EMBL: S74376; AAC60515.1; -;
 CC DR InterPro: IPR000875; Cecropin.
 CC DR InterPro: IPR003254; IIP-cecropin.
 CC DR Pfam: PF00272; cecropin; 1.
 CC DR ProDom: PD003996; IIP-cecropin; 1.
 CC DR PROSITE: PS00268; CECROPIN; 1.
 CC DR Insect Immunity: Antibiotic; Hemolymph; Amidation; Multigene family;
 CC Signal.
 CC KW SIGNAL. 1 22
 CC FT PROPEP 23 26 POTENTIAL.
 CC FT CHAIN 27 61 BY SIMILARITY.
 CC FT MOD_RES 61 61 CECROPIN A.
 CC SQ SEQUENCE 63 AA; 6762 MW; 6A1C39975510D86A CRC64;
 CC -----
 CC Query Match 38.9%; Score 42; DB 1; Length 63;
 CC Best Local Similarity 87.5%; Pred. No. 4;
 CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC OY 1 KWKLFKKT 8
 CC DB 27 KWKVFKKT 34
 CC -----
 CC RESULT 14
 CC YG31_HAEIN STANDARD; PRT; 190 AA.
 CC ID YG31_HAEIN
 CC AC P44279;
 CC DT 01-NOV-1995 (Rel. 32, Created)
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Hypothetical protein YH1631 precursor.
 CC GN H1631.
 CC OS Haemophilus influenzae.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC OC Haemophilus.
 CC OX NCBI_TaxID=727;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Rd / KW20 / ATCC 51907;
 CC RA MEDLINE-95350630; PubMed-7542800;
 CC RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 CC RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 CC RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 CC RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 CC RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 CC RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 CC RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 CC RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 CC RA Venter J.C.;
 CC RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 CC Rd.";
 CC RL Science 269:496-512(1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: U32836; AAC23283.1; -
 DR TRGR: H11631; -
 KM Hypothetical protein; Signal: Complete proteome.
 FT SIGNAL 1 28
 FT CHAIN 29 190
 FT SEQUENCE 190 AA; 22314 MW; E4E3A8C7957EC908 CRC64;
 SQ
 Query Match 38.9%; Score 42; DB 1; Length 190;
 Best Local Similarity 36.8%; Pred. No. 11;
 Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
 QY 1 KKKLFKKIGIGKFLHSAKK 19
 Db 90 EFRLFQVSMGEFLSISK 108
 RESULT 15
 KCL1_BOVIN STANDARD; PRT; 325 AA.
 AC P35506;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Casein kinase I, alpha isoform (EC 2.7.1.-) (CKI-alpha) (CKI).
 GN CSNK1A1.
 OS Bos taurus (Bovine).
 OS Oryctolagus cuniculus (Rabbit), and
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID:9913, 9986, 9940;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC SPECIES-Bovine; TISSUE-Brain;
 RX MEDLINE-92052129; PubMed-1946367;
 RA Rowles J., Slaughter C., Moomaw C., Hsu J., Cobb M.H.;
 RT Purification of casein kinase I and isolation of cDNAs encoding
 RL multiple casein kinase I-like enzymes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9548-9552(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Rabbit; TISSUE-Skeletal muscle;
 RA Zhai L., Depooll-Roach A.A., Roach P.J.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Sheep; TISSUE-Blood;
 RA Yamamoto N.;
 RT "Ovine casein kinase I alpha gene.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CASEIN KINASES ARE OPERATIONALLY DEFINED BY THEIR
 CC PREFERENTIAL UTILIZATION OF ACIDIC PROTEINS SUCH AS CASEINS
 CC AS SUBSTRATES. IT CAN PHOSPHORYLATE A LARGE NUMBER OF PROTEINS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CASEIN KINASE I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: M76543; AAA30451.1; -

DR EMBL: U59166; AAB03992.1; -
 DR EMBL: AB050945; BAB17769.1; -
 DR HSSP: Q06486; 1CKI.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR Pfam: PF00069; Ser_thr_Pkinase.
 DR ProDom: PD000001; Euk_Pkinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Multigene family.
 FT DOMAIN 17 285
 FT NP_BIND 23 31
 FT BINDING 46 46
 FT ACT_SITE 136 136
 FT SEQUENCE 325 AA; 37567 MW; B84DC84BDDC17854 CRC64;
 SQ
 Query Match 38.9%; Score 42; DB 1; Length 325;
 Best Local Similarity 61.5%; Pred. No. 19;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 KKKLFKKIGIGKFLHS 13
 Db 16 KKKLVKRGSGSF 28
 Search completed: June 6, 2003, 13:23:58
 Job time : 8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 13:21:26 ; Search time 23.5 Seconds
(without alignments)
175.359 Million cell updates/sec

Title: US-10-081-418-1
Perfect score: 108
Sequence: 1 KMKLFKKIGKIFLHSAKF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	44.4	309	16	Q8RC36 thermomaneer
2	47.5	44.0	237	17	Q9U283
3	47.5	44.0	427	16	Q9K5P7
4	47.5	44.0	501	3	Q96U53
5	47	43.5	1562	3	Q60034
6	46	42.6	376	10	Q8VYE2
7	46	42.6	414	10	Q9LG22
8	46	42.6	432	10	Q40166
9	46	42.6	666	16	Q98OC7
10	45	41.7	88	2	Q9AINE
11	45	41.7	154	11	Q9CYL5
12	45	41.7	319	5	Q9U2N8
13	45	41.7	925	16	Q8YBF0
14	44	40.7	205	16	Q99TD7
15	44	40.7	353	3	Q08980
16	44	40.7	499	5	Q9VN08

17	43	39.8	170	16	Q9RX21	Q9rx21 delinococcus
18	43	39.8	199	12	Q9IO51	Q9io51 influenzavi
19	43	39.8	199	12	Q9IO50	Q9io50 influenzavi
20	43	39.8	199	12	Q9O217	Q9o217 influenzavi
21	43	39.8	199	12	Q9O2P4	Q9o2p4 influenzavi
22	43	39.8	199	12	Q9O205	Q9o205 influenzavi
23	43	39.8	199	12	Q9O203	Q9o203 influenzavi
24	43	39.8	199	12	Q9O201	Q9o201 influenzavi
25	43	39.8	199	12	Q9O1Z9	Q9o1z9 influenzavi
26	43	39.8	199	12	Q9O1Z7	Q9o1z7 influenzavi
27	43	39.8	199	12	Q9O1Z5	Q9o1z5 influenzavi
28	43	39.8	199	12	Q9O1Z3	Q9o1z3 influenzavi
29	43	39.8	199	12	Q9O1Z1	Q9o1z1 influenzavi
30	43	39.8	199	12	Q9O1Y9	Q9o1y9 influenzavi
31	43	39.8	199	12	Q9O1Y7	Q9o1y7 influenzavi
32	43	39.8	199	12	Q8QZ11	Q8qz11 influenzavi
33	43	39.8	199	12	Q8QZ10	Q8qz10 influenzavi
34	43	39.8	199	12	Q8QZ09	Q8qz09 influenzavi
35	43	39.8	199	12	Q8QZ08	Q8qz08 influenzavi
36	43	39.8	199	12	Q8QZ07	Q8qz07 influenzavi
37	43	39.8	199	12	Q8QZ06	Q8qz06 influenzavi
38	43	39.8	199	12	Q8QZ05	Q8qz05 influenzavi
39	43	39.8	199	12	Q8QZ04	Q8qz04 influenzavi
40	43	39.8	199	12	Q8QZ03	Q8qz03 influenzavi
41	43	39.8	199	12	Q8QZ02	Q8qz02 influenzavi
42	43	39.8	199	12	Q8QZ01	Q8qz01 influenzavi
43	43	39.8	199	12	Q8QZ00	Q8qz00 influenzavi
44	43	39.8	199	12	Q8QZ00	Q8qz00 influenzavi
45	43	39.8	199	12	Q8QZ00	Q8qz00 influenzavi

ALIGNMENTS

RESULT 1				
ID	Q8RC36	PRELIMINARY;	PRT;	309 AA.
AC	Q8RC36;			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components.			
GN	DPPB4 OR TTE0612.			
OS	Thermomaneerobacter tengcongensis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;			
OC	Thermomaneerobacteriales; Thermomaneerobacteriaceae; Thermomaneerobacter.			
OX	NCBI_TaxID=119072;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-MB4T / JCM11007;			
RX	MEDLINE=21992816; PubMed=11997336;			
RA	Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,			
RA	Tan H., Chen R., Wang J., Yu J., Yang H.;			
RT	"A complete sequence of T. tengcongensis genome.";			
RL	Genome Res. 12:689-700(2002).			
DR	EMBL; AE013030; AAM23882.1; -			
KW	Complete proteome.			
SQ	SEQUENCE 309 AA; 34754 MW; A90E0686091EFF74 CRC64;			
Query Match		44.4%;	Score 48;	DB 16; Length 309;
Best Local Similarity		56.2%;	Pred. No. 11;	
Matches	9; Conservative	1; Mismatches	6; Indels	0; Gaps
QY	1 KMKLFKKIGKIFLHS 16			
DB	157 KMKLFPPVSGWGTFAHT 172			
QY				
RESULT 2				
ID	Q9U283	PRELIMINARY;	PRT;	237 AA.

AC Q9U283;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Arginase related (EC 3.5.3.1).
 GN PAB1359.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ORSAY;
 RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 structure and evolution."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ248287; CAB50176.1; -.
 DR InterPro: IPR000287; Arginase.
 DR Pfam: PF00491; arginase; 1.
 DR PROSITE: PS00148; ARGINASE_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 237 AA; 27191 MW; E044EF7EB136463D CRC64;

Query Match 44.0%; Score 47.5; DB 17; Length 237;
 Best Local Similarity 39.1%; Pred. No. 9.9;
 Matches 9; Conservative 6; Mismatches 5; Indels 3; Gaps 1;

OY 1 KMKLF---KKIGGKFLHSAKKF 20
 DB 135 KMKAYINLKOHGKRVIREARKF 157

RESULT 3

O9K5P7 PRELIMINARY; PRT; 427 AA.
 AC O9K5P7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein BH4041.
 GN BH4041.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed-11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001520; BAB07760.1; -.
 DR InterPro: IPR003346; Transposase_20.
 DR InterPro: IPR002525; Transposase_9.
 DR Pfam: PF02371; Transposase_20; 1.
 DR Pfam: PF01548; Transposase_9; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 427 AA; 48695 MW; F7E8099BCDA45C5C CRC64;

Query Match 44.0%; Score 47.5; DB 16; Length 427;
 Best Local Similarity 45.8%; Pred. No. 18;
 Matches 11; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

OY 1 KMKLFKKIGIG----KFLHSARK 19
 DB 217 KMKPFVKRGVIGIKRATKLVETAKK 240

RESULT 4

O96U53 PRELIMINARY; PRT; 501 AA.
 AC O96U53;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 56.9 kDa protein.
 GN B24P7.050.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohnsels J., Brandt P., Farlmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL389890; CAD11797.1; -.
 DR InterPro: IPR000822; ZnF_C2H2.
 DR SMART: SM00355; ZnF_C2H2; 3.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_2.
 KW DNA-binding; Hypothetical protein; Zinc-finger.
 SQ SEQUENCE 501 AA; 56866 MW; 5B7D09A2ED6C90E9 CRC64;

Query Match 44.0%; Score 47.5; DB 3; Length 501;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

OY 2 WKLFK-KIGGKFLHSAKKF 20
 DB 320 WTLMKPCSGCHKYHSAKFF 339

RESULT 5

O60034 PRELIMINARY; PRT; 1562 AA.
 AC O60034;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ATP-binding multidrug cassette transporter.
 OS Botrytis cinerea (Botryotinia fuckeliana).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
 OC Helotiales; Sclerotiniaceae; Botryotinia.
 OX NCBI_TaxID=40559;
 RN [1]
 RP SEQUENCE OF 1-1283 FROM N.A.
 RC STRAIN=SAS56;
 RA Del Sorbo G., van Kan J.A.L., De Waard M.A.;

RT "Characterization of a P-glycoprotein-like multidrug resistance gene
 in the plant pathogenic fungus Botrytis cinerea."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-1283 FROM N.A.
 RC STRAIN=SAS56;
 RX MEDLINE=97324590; PubMed-9180695;
 RA Del Sorbo G., Van Nistelrooy H., Andrade A., van Kan J.A.L., Balzi,
 RA De Waard M.;

RT "Multidrug resistance in Aspergillus nidulans involves novel ATP-
 binding cassette transporters."
 RL Mol. Gen. Genet. 254:417-426(1997).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: Z68906; CA93142.1; -.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR000577; FGGY_kin.
 DR InterPro: IPR001092; HLH_basic.
 DR Pfam: PF00005; ABC_tran; 2.

DR ProDom; PD000006; ABC_transport; 2.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR PROSITE; PS00445; PGCG_KINASES_2; UNKNOWN_1.
 DR PROSITE; PS0038; HELIX_LOOP_HELIX; UNKNOWN_1.
 KW ATP-binding; transport.
 SQ SEQUENCE 1562 AA; 173535 MW; 89036ECCE3D24919 CRC64;

Query Match 43.5%; Score 47; DB 3; Length 1562;
 Best Local Similarity 30.8%; Pred. No. 82;
 Matches 8; Conservative 7; Mismatches 3; Indels 8; Gaps 1;

OY 1 KWK-----LFKKIGIGKFLHSAKF 18
 DB 1497 KWRNFGIFLIGLYOQLCVGLFLHQR.1522

RESULT 6
 O8VEE2 PRELIMINARY; PRT; 376 AA.
 ID O8VEE2
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 42.6 kDa protein.
 GN AIG55900
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Ouach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carinanci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT "Full Length cDNA of gene AIG55900 (GI:15222797)."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AV072140; AAL59962.1; -
 DR InterPro; IPR004274; NIF.
 DR Pfam; PF03031; NIF; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 376 AA; 42639 MW; 160466DD597215AA CRC64;

Query Match 42.6%; Score 46; DB 10; Length 376;
 Best Local Similarity 42.1%; Pred. No. 28;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 2 WKLFFKIGIGKFLHSAKF 20
 DB 212 WRTFKRPGVDALFHLGKF 230

RESULT 7
 O9LG22 PRELIMINARY; PRT; 414 AA.
 ID O9LG22
 AC O9LG22
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE F14316.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.

RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
 RA Kim C., Altafi H., Bel O., Chin C., Chou J., Choi E., Conn L.,
 RA Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
 RA Lanz C., Li J., Liu A., Liu K., Liu S., Mukharasy N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaverl A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F14316 from chromosome
 I.";
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Khan S., Kim C., Altafi H., Bel O., Chin C., Chou J., Choi E.,
 RA Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaverl A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC002304; AAF79316.1; -
 DR InterPro; IPR004274; NIF.
 DR Pfam; PF03031; NIF; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 414 AA; 46605 MW; 1D21A52D01F8D413 CRC64;

Query Match 42.6%; Score 46; DB 10; Length 414;
 Best Local Similarity 42.1%; Pred. No. 31;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 2 WKLFFKIGIGKFLHSAKF 20
 DB 212 WRTFKRPGVDALFHLGKF 230

RESULT 8
 ID O40166 PRELIMINARY; PRT; 432 AA.
 AC O40166
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Mutant phytoene synthase.
 GN G70M5.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-ATISA CRAIG, TISSUE-PERICARP;
 RX MEDLINE=93344508; PubMed=8343597;
 RA Fray R.G., Giersen D.;
 RT "Identification and genetic analysis of normal and mutant phytoene
 RT synthase genes by of tomato by sequencing, complementation and co-
 RT suppression.";
 RL Plant Mol. Biol. 22:589-602(1993).

RESULT 12

0902N8 PRELIMINARY; PRT; 319 AA.
 AC 0902N8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE Y37H2A.8 protein.
 GN Y37H2A.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA White S.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;

RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: AL132863; CAB60572.1;
 DR InterPro: IPR002900; DUF38.
 DR InterPro: IPR001810; F-box.
 DR Pfam: PF01827; DUF38; 1.
 DR Pfam: PF00646; F-box; 1.
 SQ SEQUENCE 319 AA; 37863 MW; 69C3C1E2CEA8FD14 CRC64;

Query Match 41.7%; Score 45; DB 5; Length 319;
 Best Local Similarity 46.7%; Pred. No. 35;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 OY 1 KMKLFKKIGIGKFLH 15
 Db 249 EMMKSKDTKLGKFFH 263

RESULT 13

08YBFO PRELIMINARY; PRT; 925 AA.
 AC 08YBFO;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Nitrate reductase alpha chain (BC 1.7.99.4).
 GN BME110950.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kaparat I., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhatnagar A., Lykakis A., Reznik G.,
 RA Jablonki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
 RA Selkov E., Elizer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Kasselorn R., Kyprides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009728; AL54192.1;
 DR InterPro: IPR001467; Prok_Moxred.
 DR Pfam: PF00384; molybdopterin; 1.
 DR Pfam: PF01568; molybdopterin; 1.
 DR PROSITE: PS00490; MOLYBDOPTERIN_PROK_2; 1.
 DR PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; 1.
 KW Oxidoreductase; Complete proteome
 SQ SEQUENCE 925 AA; 102574 MW; 15278350CB06033F CRC64;

Query Match 41.7%; Score 45; DB 16; Length 925;
 Best Local Similarity 50.0%; Pred. No. 1e+02;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 KMKLFKKIGIGKFLHSAKF 20
 Db 72 KMKLEKDGKDKVELAKSF 91

RESULT 14

099TD7 PRELIMINARY; PRT; 205 AA.
 AC 099TD7;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein SAV1727.
 GN SAV1727 OR SA1548.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=158878, 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cul L., Oguchi A., Aoki K.-I., Nagai Y., Itoh T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus."
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AP003363; BAB57889.1;
 DR EMBL: AP003134; BAB42816.1;
 DR InterPro: IPR004552; Acyltransferase.
 DR Pfam: PF01553; Acyltransferase; 1.
 DR TIGRFAMs: TIGR00530; AGP_acyltn; 1.
 KW Hypothetical protein; Complete proteome
 SQ SEQUENCE 205 AA; 23074 MW; 0078A4DC566218E CRC64;

Query Match 40.7%; Score 44; DB 16; Length 205;
 Best Local Similarity 55.0%; Pred. No. 32;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 KMKLFKKIGIGKFLHSAKF 20
 Db 63 KMKLFKKIGIGKFLHSAKF 82

RESULT 15

008980 PRELIMINARY; PRT; 353 AA.
 ID 008980;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Chromosome XVI reading frame ORF YP1264C.
 GN YP1264C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Delius H., Hebling U.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RA Duysterhoert A., Floeth M., Filtz M., Hilbert H., Moestl D.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z73620; CA97999.1; -.
DR SGD: S0006185; YPL264C.
DR InterPro: IPR000620; DUF6.
DR Pfam: PF00892; DUF6; 2.
SQ SEQUENCE 353 AA; 39702 MW; 68A971BC5DB92988 CRC64;

Query Match 40.7%; Score 44; DB 3; Length 353;
Best Local Similarity 46.7%; Pred. No. 56;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KWKLFRKIGIGKFLH 15
: | | | : | | | : | | | : | | |
Db 253 QWGLFLNLGIGSGFIH 267

Search completed: June 6, 2003, 13:24:53
Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 13:22:41 ; Search time 15.5 seconds

(Without alignments)
133.213 Million cell updates/sec

Title: US-10-081-418-1

Perfect score: 108

Sequence: 1 KMKLFKRGIGKFLHSAKRF 20

Scoring table: BLOSUM62

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

Published Applications AA: *
1: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	20	US-10-081-418-1	Sequence 1, Appli
2	64	59.3	32	US-09-908-139-6	Sequence 6, Appli
3	63	58.3	23	US-09-820-053A-7	Sequence 7, Appli
4	63	58.3	23	US-09-904-753-2	Sequence 2, Appli
5	63	58.3	23	US-09-030-619-211	Sequence 21, App
6	63	58.3	26	US-09-904-753-13	Sequence 13, Appl
7	63	58.3	26	US-09-030-619-120	Sequence 120 App
8	63	58.3	28	US-09-908-139-1	Sequence 1, Appli
9	63	58.3	28	US-09-909-652-7	Sequence 121, App
10	63	58.3	28	US-09-030-619-121	Sequence 4, Appli
11	63	58.3	303	US-09-917-340-4	Sequence 146, App
12	60	55.6	23	US-09-820-053A-146	Sequence 9, Appli
13	59	54.6	33	US-09-908-139-9	Sequence 9, Appli
14	56	51.9	23	US-09-904-753-1	Sequence 210, App
15	56	51.9	23	US-09-030-619-210	Sequence 3, Appli
16	54	50.0	22	US-09-807-720-3	Sequence 33, Appl
17	51	47.2	15	US-09-820-053A-33	Sequence 24, Appl
18	50	46.3	22	US-09-820-053A-24	Sequence 3, Appli
19	50	46.3	22	US-09-904-753-3	

20	50	46.3	22	US-09-904-753-4	Sequence 4, Appli
21	49.5	45.8	31	US-09-908-139-8	Sequence 8, Appli
22	48.5	44.9	23	US-09-908-139-12	Sequence 12, Appl
23	48	44.4	30	US-09-908-139-7	Sequence 7, Appli
24	48	44.4	376	US-09-738-626-6975	Sequence 6975, Ap
25	46.5	43.1	30	US-09-908-139-11	Sequence 11, Appl
26	46	42.6	412	US-09-371-307-76	Sequence 76, Appl
27	45	41.7	37	US-09-904-753-5	Sequence 5, Appli
28	45	41.7	37	US-09-030-619-176	Sequence 176, App
29	45	41.7	535	US-10-020-079-16	Sequence 16, Appl
30	45	41.7	548	US-10-020-079-12	Sequence 12, Appl
31	45	41.7	648	US-10-020-079-14	Sequence 14, Appl
32	45	41.7	654	US-10-020-079-10	Sequence 10, Appl
33	45	41.7	751	US-10-020-079-8	Sequence 8, Appli
34	45	41.7	764	US-10-020-079-6	Sequence 6, Appli
35	45	41.7	776	US-10-020-079-24	Sequence 24, Appl
36	45	41.7	789	US-10-020-079-22	Sequence 22, Appl
37	45	41.7	838	US-10-020-079-40	Sequence 40, Appl
38	45	41.7	851	US-10-020-079-38	Sequence 38, Appl
39	45	41.7	863	US-10-020-079-32	Sequence 32, Appl
40	45	41.7	864	US-10-020-079-4	Sequence 4, Appli
41	45	41.7	870	US-10-020-079-2	Sequence 2, Appli
42	45	41.7	876	US-10-020-079-30	Sequence 30, Appl
43	45	41.7	889	US-10-020-079-18	Sequence 18, Appl
44	45	41.7	895	US-10-020-079-10	Sequence 10, Appl
45	45	41.7	951	US-10-020-079-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-10-081-418-1
Sequence 1, Application US/10081418,
Publication No. US20030096745A1
GENERAL INFORMATION:
APPLICANT: HAHM, Kyung-Soo
APPLICANT: LEE, Yong Gun
APPLICANT: PARK, Yeonkyung
TITLE OF INVENTION: No. US20030096745A1 peptides with increased + charge and hyd.
TITLE OF INVENTION: substituting one or more amino acids of CA-MA peptide and
FILE REFERENCE: 428.1014
CURRENT APPLICATION NUMBER: US/10/081,418
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Knapten 1.71
SEQ ID NO 1
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: CA-MA peptide made by fusing 1-8 amino acid of secretin A and
US-10-081-418-1
Query Match
Best local similarity 100.0%; Score 108; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KMKLFKRGIGKFLHSAKRF 20
Db 1 KMKLFKRGIGKFLHSAKRF 20
RESULT 2
US-09-908-139-6
Sequence 6, Application US/09908139
Publication No. US20030096949A1
GENERAL INFORMATION:
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha A.

APPLICANT: Patrzykat, Aleksander
APPLICANT: Woods, Donald
APPLICANT: Jia, Xiaoyan
TITLE OF INVENTION: ANTI-ENDOTOXIC, ANTIMICROBIAL CATIONIC
FILE OF INVENTION: PEPTIDES AND METHODS OF USE THEREFOR
FILE REFERENCE: 07422/016001
CURRENT APPLICATION NUMBER: US/09/908,139
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: 09/143,124
PRIOR FILING DATE: 1998-08-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 32
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: antimicrobial cationic peptide
US-09-908-139-6

Query Match 59.3%; Score 64; DB 9; Length 32;
Best Local Similarity 72.2%; Pred. No. 0.0017;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KWKLFKRGIGKFLHSAK 18
DB 1 KWKLFKRGIGAVLKVK 18

RESULT 3
US-09-820-053A-7
Sequence 7, Application US/09820053A
Publication No. US20030083243A1
GENERAL INFORMATION:
APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
FILE REFERENCE: HELX027
CURRENT APPLICATION NUMBER: US/09/820,053A
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 165
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 23
TYPE: PRT
ORGANISM: ARTIFICIAL SEQUENCE
FEATURE:
OTHER INFORMATION: SYNTHETIC SEQUENCE
NAME/KEY: MOD_RES
LOCATION: (23)
OTHER INFORMATION: AMIDATION
US-09-820-053A-7

Query Match 58.3%; Score 63; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 GIGKFLHSAKFF 20
DB 1 GIGKFLHSAKFF 12

RESULT 4
US-09-904-753-2
Sequence 2, Application US/09904753
Publication No. US20030092612A1
GENERAL INFORMATION:
APPLICANT: Lynos, Robert T
TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
FILE REFERENCE: 2973 ver 2
CURRENT APPLICATION NUMBER: US/09/904,753
CURRENT FILING DATE: 2001-07-13

PRIOR APPLICATION NUMBER: WO 96/25183
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 23
TYPE: PRT
ORGANISM: Xenopus laevis
US-09-904-753-2

Query Match 58.3%; Score 63; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 GIGKFLHSAKFF 20
DB 1 GIGKFLHSAKFF 12

RESULT 5
US-09-030-619-211
Sequence 21, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Ertle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INJECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 211
LENGTH: 23
TYPE: PRT
ORGANISM: Xenopus laevis
US-09-030-619-211

Query Match 58.3%; Score 63; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 GIGKFLHSAKFF 20
DB 1 GIGKFLHSAKFF 12

RESULT 6
US-09-904-753-13
Sequence 13, Application US/09904753
Publication No. US20030092612A1
GENERAL INFORMATION:
APPLICANT: Lynos, Robert T
TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
FILE REFERENCE: 2973 ver 2
CURRENT APPLICATION NUMBER: US/09/904,753
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: WO 96/25183
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (26)
OTHER INFORMATION: Xaa at position 26 is Ser-amide
OTHER INFORMATION: Description of Artificial Sequence: Hybrid
OTHER INFORMATION: antimicrobial peptide
US-09-904-753-13

Query Match 58.3%; Score 63; DB 9; Length 26;
Best Local Similarity 85.7%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKLFKKIGIGKFL 14
Db 1 KWKLFKKIGIGAVL 14

RESULT 7
US-09-030-619-120
Sequence 120, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 120
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-120

Query Match 58.3%; Score 63; DB 10; Length 26;
Best Local Similarity 85.7%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKLFKKIGIGKFL 14
Db 1 KWKLFKKIGIGAVL 14

RESULT 8
US-09-908-139-1
Sequence 1, Application US/09908139
Publication No. US20030096949A1
GENERAL INFORMATION:
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha A.
APPLICANT: Patrzykat, Aleksander
APPLICANT: Woods, Donald
APPLICANT: Jia, Xiaoyan
TITLE OF INVENTION: ANTI-ENDOTOXIC, ANTIMICROBIAL CATIONIC
TITLE OF INVENTION: PEPTIDES AND METHODS OF USE THEREFOR
FILE REFERENCE: 07422/016001
CURRENT APPLICATION NUMBER: US/09/908,139
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: 09/143,124
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1

LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: antimicrobial cationic peptide
US-09-908-139-1

Query Match 58.3%; Score 63; DB 9; Length 28;
Best Local Similarity 85.7%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKLFKKIGIGKFL 14
Db 1 KWKLFKKIGIGAVL 14

RESULT 9
US-09-909-652-7
Sequence 7, Application US/09909652
Patent No. US20020025537A1
GENERAL INFORMATION:
APPLICANT: Kairos Scientific, Inc.
APPLICANT: Bylina, Edward J.
APPLICANT: Coleman, William J.
APPLICANT: Youvan, Douglas C.
TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR GENERATING
TITLE OF INVENTION: AND SCREENING COMPOUNDS THAT AFFECT CELL VIABILITY
FILE REFERENCE: 22346-7001
CURRENT APPLICATION NUMBER: US/09/909,652
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/219,179
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: CEMA analogue of ceceopin-melittin hybrid
US-09-909-652-7

Query Match 58.3%; Score 63; DB 10; Length 28;
Best Local Similarity 85.7%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKLFKKIGIGKFL 14
Db 1 KWKLFKKIGIGAVL 14

RESULT 10
US-09-030-619-121
Sequence 121, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 121
LENGTH: 28
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-121

Query Match 58.3%; Score 63; DB 10; Length 28;
Best Local Similarity 85.7%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMKLFKIGIGKFL 14
DB 1 KMKLFKIGIGAVL 14

RESULT 11
US-09-917-340-4
Sequence 4, Application US/09917340
Patent No. US20020090369A1
GENERAL INFORMATION:
APPLICANT: Murphy, Christopher J.
APPLICANT: McAnulty, Jonathan F.
APPLICANT: Reid, Ted W.
TITLE OF INVENTION: Transplant Media
FILE REFERENCE: TPLANT-06468
CURRENT APPLICATION NUMBER: US/09/917,340
CURRENT FILING DATE: 2001-07-29
PRIOR APPLICATION NUMBER: 60/221,632
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/249,602
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/290,932
PRIOR FILING DATE: 2001-05-15
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 303
TYPE: PRT
ORGANISM: Xenopus laevis
US-09-917-340-4

Query Match 58.3%; Score 63; DB 10; Length 303;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GIGKFLSAKFF 20
DB 83 GIGKFLSAKFF 94

RESULT 12
US-09-820-053A-146
Sequence 146, Application US/09820053A
Publication No. US20030083243A1
GENERAL INFORMATION:
APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
FILE REFERENCE: HELIX027
CURRENT APPLICATION NUMBER: US/09/820,053A
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 165
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 146
LENGTH: 23
TYPE: PRT
ORGANISM: ARTIFICIAL SEQUENCE
FEATURE:
OTHER INFORMATION: SYNTHETIC SEQUENCE
NAME/KEY: MOD_RES
LOCATION: (23)
OTHER INFORMATION: AMIDATION
US-09-820-053A-146

Query Match 55.6%; Score 60; DB 9; Length 23;

Best Local Similarity 91.7%; Pred. No. 0.0054;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 GIGKFLSAKFF 20
DB 1 GIGKFLSAKFF 12

RESULT 13
US-09-908-139-9
Sequence 9, Application US/09908139
Publication No. US20030096949A1
GENERAL INFORMATION:
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha A.
APPLICANT: Patrzykat, Aleksander
APPLICANT: Woods, Donald
APPLICANT: Jia, Xiaoyan
TITLE OF INVENTION: ANTI-ENDOTOXIC, ANTIMICROBIAL CATIONIC
PEPTIDES AND METHODS OF USE THEREFOR
FILE REFERENCE: 07422/016001
CURRENT APPLICATION NUMBER: US/09/908,139
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: 09/143,124
PRIOR FILING DATE: 1998-08-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 33
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: antimicrobial cationic peptide
US-09-908-139-9

Query Match 54.6%; Score 59; DB 9; Length 33;
Best Local Similarity 70.6%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 WKLFKIGIGKFLSAK 18
DB 3 WKLFKIGIGAVLKVLK 19

RESULT 14
US-09-904-753-1
Sequence 1, Application US/09904753
Publication No. US20030092612A1
GENERAL INFORMATION:
APPLICANT: Lynos, Robert T.
TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
Ophthalmic Preparations, Including Solutions,
Emulsions, and Suspensions
FILE REFERENCE: 2973 ver 2
CURRENT APPLICATION NUMBER: US/09/904,753
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: WO 96/25183
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 23
TYPE: PRT
ORGANISM: Xenopus laevis
PUBLICATION INFORMATION:
AUTHORS: Lee et al.
TITLE: High-Level Expression of Antimicrobial Peptide Mediated
TITLE: by a Fusion Partner Reinforcing Formation of Inclusion
Bodies
JOURNAL: Biochem. Biophys. Res. Commun.
VOLUME: 277
PAGES: 575-580
DATE: Sept 21, 2000

US-09-904-753-1

Query Match 51.9%; Score 56; DB 9; Length 23;
 Best Local Similarity 91.7%; Pred. No. 0.024;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 GIGKFLHSAGKF 20
 |||||
 Db 1 GIGKFLHSAGKF 12

RESULT 15

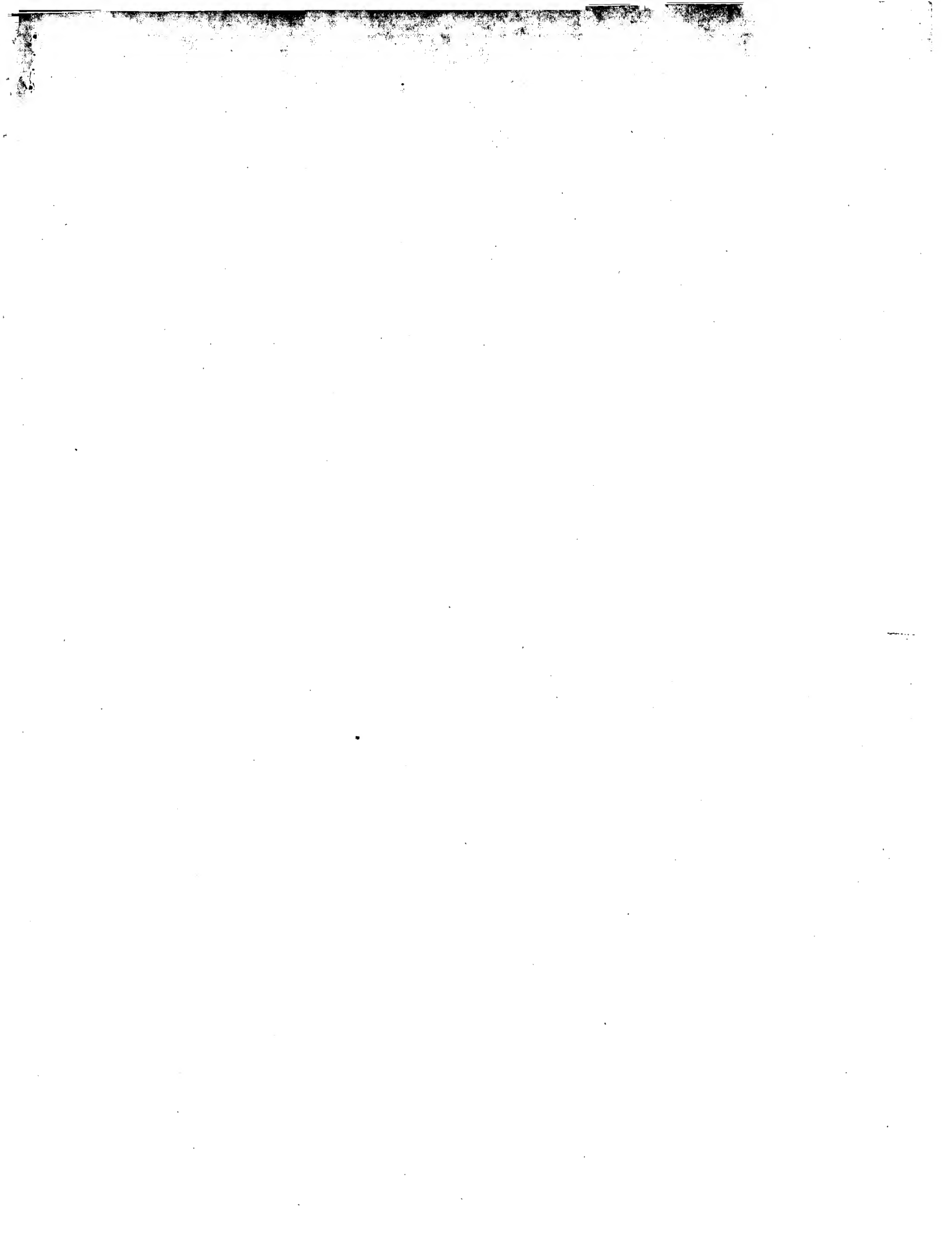
US-09-030-619-210
 ; Sequence 210, Application US/09030619B
 ; Patent No. US20020035061A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krieger, Timothy J.
 ; APPLICANT: Taylor, Robert
 ; APPLICANT: Erile, Douglas
 ; APPLICANT: Fraser, Janet R.
 ; APPLICANT: West, Michael H.P.
 ; APPLICANT: McNicol, Patricia J.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
 ; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
 ; FILE REFERENCE: 660081.406
 ; CURRENT APPLICATION NUMBER: US/09/030,619B
 ; CURRENT FILING DATE: 1998-02-25
 ; NUMBER OF SEQ ID NOS: 232
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 210
 ; LENGTH: 23
 ; TYPE: PRT
 ; ORGANISM: Xenopus laevis
 US-09-030-619-210

Query Match 51.9%; Score 56; DB 10; Length 23;
 Best Local Similarity 91.7%; Pred. No. 0.024;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 GIGKFLHSAGKF 20
 |||||
 Db 1 GIGKFLHSAGKF 12

Search completed: June 6, 2003, 13:26:02
 Job time : 15.5 secs

3 Burkhead
 Tylenol



[illegible]

PA (AMGE-) AMGEN INC.

DR WPI; 1992-064700/08

PA (AMGE-) AMGEN INC.

DR WPI; 1992-064700/08

PT Method for inhibiting target cell and virus growth - comprises
 PT administering amphiphilic peptide compns, useful for treating
 PT viral and phytopathogenic infections, tumours and burns
 XX
 PS Disclosure; Page 60; 72pp; English.
 XX
 CC The peptides of the invention are effective pharmaceuticals having
 CC anti-microbial, anti-viral and anti-tumour activity. They are also
 CC useful for inhibiting, preventing or destroying the motility of
 CC sperm and hence have application in a spermicide preparation. They
 CC also have anti-parasitic activity and are useful in wound healing.
 CC as preservatives and sterilants and to inhibit growth of
 CC phytopathogenic microorganisms. AAR20969 and AAR20970 were published in
 CC Haighten and Ostresh, Bio Chromatography, Vol 2, Issue 2, page 80-83,
 CC 1987.
 XX
 SQ Sequence 30 AA;
 QY
 Query Match 61.0%; Score 64; DB 13; Length 30;
 Best Local Similarity 80.0%; Pred. No. 0.12;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Db 1 KKKKLLKKPPPLKKLLKKLL 20
 4 KKKKLLKKLLKKLLKKLLKKLL 23
 RESULT 4
 AAR22883
 ID AAR22883 standard; Peptide; 30 AA.
 XX AAR22883;
 AC
 DT 22-AUG-1992 (first entry)
 XX
 DE Amphiphilic peptide to inhibit growth of a target cell.
 XX
 KW Antimicrobial; antifungal; antitumour; pharmaceuticals; antibiotics;
 KW antiparasitic; spermicides; burns; wound healing.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site /note="acylated"
 FT Modified-site 30
 FT Modified-site /note="amidated"
 XX
 PN CA2047317-A.
 PD 20-JAN-1992.
 XX
 PF 18-JUL-1991; 91CA-2024317.
 XX
 PR 08-JUL-1991; 91US-0725331.
 PR 19-JUL-1990; 90US-0554442.
 XX
 PA (SCRI-) SCRIPPS RES INST.
 PI Houghten RA, Blondelle S;
 PI Houghten RA, Blondelle S;
 DR WPI; 1992-114943/15.
 XX
 XX Amphiphilic peptide(s) and analogues - for use in e.g.
 PT antimicrobial, antifungal or antitumour compositions, having
 PT increased biological activity
 XX
 PS Claim 30; Page 60; 71pp; English.
 XX
 CC The amphiphilic peptide (SEQ ID NO 62) was prepd. by standard
 CC solid phase methods. The peptide is an analogue of the peptide
 CC having SEQ ID NO 3 (AAR22824), with an N-terminal extension of 12
 CC amino acids. Substitution and deletion analogues of this peptide

CC have increased biological activity and are effective as pharma-
 CC ceuticals e.g. antibiotics for bacterial, fungal or viral
 CC infections, or in spermicides or antitumour or antiparasitic
 CC agents. Additionally the peptides can be used in wound healing
 CC compns. or for treating burns or other skin or eye infections.
 CC See also AAR2822-89.
 XX
 SQ Sequence 30 AA;
 QY
 Query Match 61.0%; Score 64; DB 13; Length 30;
 Best Local Similarity 80.0%; Pred. No. 0.12;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Db 1 KKKKLLKKPPPLKKLLKKLL 20
 4 KKKKLLKKLLKKLLKKLLKKLL 23
 RESULT 5
 AAR21400
 ID AAR21400 standard; peptide; 36 AA.
 XX AAR21400;
 AC
 DT 16-MAY-1992 (first entry)
 XX
 DE Sequence of amphiphilic peptide SEQ ID NO. 63 with
 DE acetylated N-terminus, may be a C-terminal amide.
 XX
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;
 KW spermicide; wound healing; sterilant.
 XX
 PN WO9201462-A.
 PD 06-FEB-1992.
 XX
 PF 17-JUL-1991; 91WO-US05047.
 XX
 PR 19-JUL-1990; 90US-0554422.
 XX
 PA (SCRI-) SCRIPPS RES INST.
 PI Houghten RA, Blondelle S;
 PI Houghten RA, Blondelle S;
 DR WPI; 1992-064700/08.
 XX
 PT Method for inhibiting target cell and virus growth - comprises
 PT administering amphiphilic peptide compns, useful for treating
 PT viral and phytopathogenic infections, tumours and burns
 XX
 PS Disclosure; Page 61; 72pp; English.
 XX
 CC The peptides of the invention are effective pharmaceuticals having
 CC anti-microbial, anti-viral and anti-tumour activity. They are also
 CC useful for inhibiting, preventing or destroying the motility of
 CC sperm and hence have application in a spermicide preparation. They
 CC also have anti-parasitic activity and are useful in wound healing,
 CC as preservatives and sterilants and to inhibit growth of
 CC phytopathogenic microorganisms. AAR20969 and AAR20970 were published in
 CC Haighten and Ostresh, Bio Chromatography, Vol 2, Issue 2, page 80-83,
 CC 1987.
 XX
 SQ Sequence 36 AA;
 QY
 Query Match 61.0%; Score 64; DB 13; Length 36;
 Best Local Similarity 80.0%; Pred. No. 0.14;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Db 1 KKKKLLKKPPPLKKLLKKLL 20
 10 KKKKLLKKLLKKLLKKLLKKLL 29

RESULT 6
 AAR22884 ID AAR22884 standard; Peptide; 36 AA.
 XX
 AC AAR22884;
 XX
 XX 22-AUG-1992 (first entry)
 DE Amphiphilic peptide to inhibit growth of a target cell.
 XX
 KW Antimicrobial; antifungal; antitumour; pharmaceuticals; antibiotics;
 KM antiparasitic; spermicides; burns; wound healing.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note- "acylated"
 FT Modified-site 36 /note- "amidated"
 FT
 XX CA2047317-A.
 PN 20-JAN-1992.
 XX
 PD 18-JUL-1991; 91CA-2024317.
 XX
 PF 08-JUL-1991; 91US-0725331.
 XX
 PR 19-JUL-1990; 90US-0554442.
 XX
 XX (SCRI-) SCRIPPS RES INST.
 PA Houghten RA, Blondelle S;
 PI WPI; 1992-114943/15.
 DR
 XX Amphiphilic peptide(s) and analogues - for use in e.g.
 PT antimicrobial, antifungal or antitumour compositions, having
 PT increased biological activity
 XX
 XX Claim 30; Page 61; 71pp; English.
 XX
 CC The amphiphilic peptide (SEQ ID NO 63) was prepd. by standard
 CC solid phase methods. The peptide is an analogue of the peptide
 CC having SEQ ID NO 3 (AAR22824), with an N-terminal extension of 18
 CC amino acids. Substitution and deletion analogues of this peptide
 CC have increased biological activity and are effective as pharma-
 CC ceuticals e.g. antibiotics for bacterial, fungal or viral
 CC infections, or in spermicides or antitumour or antiparasitic
 CC agents. Additionally the peptides can be used in wound healing
 CC compns. or for treating burns or other skin or eye infections.
 CC See also AAR22822-89.
 CC
 XX
 SQ Sequence 36 AA;
 QY Query Match 61.0%; Score 64; DB 13; Length 36;
 Best Local Similarity 80.0%; Pred. No. 0.14;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 DB 1 KKKLKKPPPLKKLKKL 20
 10 KKKLKKLKKLKKLKKLKKL 29
 RESULT 7
 AAR21398 ID AAR21398 standard; peptide; 26 AA.
 XX
 AC AAR21398;
 XX
 XX 16-MAY-1992 (first entry)
 DE Sequence of amphiphilic peptide SEQ ID NO. 61 with

DE acetylated N-terminus, may be a C-terminal amide.
 XX
 XX Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;
 KW spermicide; wound healing; steriliant.
 XX
 XX WO9201462-A.
 PN 06-FEB-1992.
 XX
 PD 17-JUL-1991; 91WO-US05047.
 XX
 PF 19-JUL-1990; 90US-0554422.
 XX
 PR (SCRI-) SCRIPPS RES INST.
 XX
 PA Houghten RA, Blondelle S;
 PI WPI; 1992-064700/08.
 DR
 XX Method for inhibiting target cell and virus growth - comprises
 PT administering amphiphilic peptide compns, useful for treating
 PT viral and phytopathogenic infections, tumours and burns
 XX
 PS Disclosure; Page 60; 72pp; English.
 XX
 CC The peptides of the invention are effective pharmaceuticals having
 CC anti-microbial, anti-viral and anti-tumour activity. They are also
 CC useful for inhibiting, preventing or destroying the mobility of
 CC sperm and hence have application in a spermicide preparation. They
 CC also have anti-parasitic activity and are useful in wound healing,
 CC as preservatives and sterilants and to inhibit growth of
 CC phytopathogenic microorganisms. AAR20969 and AAR20970 were published in
 CC Houghten and Ostresh, Bio Chromatography, Vol 2, Issue 2, page 80-83,
 CC 1987.
 CC
 XX
 SQ Sequence 26 AA;
 QY Query Match 58.1%; Score 61; DB 13; Length 26;
 Best Local Similarity 83.3%; Pred. No. 0.25; 3; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 3;
 DB 3 KKKLKKPPPLKKLKKL 20
 2 KKKLKKLKKLKKLKKLKKL 19
 RESULT 8
 AAR22882 ID AAR22882 standard; Peptide; 26 AA.
 XX
 AC AAR22882;
 XX
 XX 22-AUG-1992 (first entry)
 DE Amphiphilic peptide to inhibit growth of a target cell.
 XX
 KW Antimicrobial; antifungal; antitumour; pharmaceuticals; antibiotics;
 KM antiparasitic; spermicides; burns; wound healing.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note- "acylated"
 FT Modified-site 26 /note- "amidated"
 FT
 XX CA2047317-A.
 PN 20-JAN-1992.
 XX
 PD 18-JUL-1991; 91CA-2024317.
 XX
 PF

PR 08-JUL-1991; 91US-0725331.
 PR 19-JUL-1990; 90US-0554442.
 PA (SCRI-) SCRIPPS RES INST.
 XX
 PI Houghten RA, Blondelle S;
 DR WPI; 1992-114943/15.
 XX
 XX Amphiphilic peptide(s) and analogues - for use in e.g.
 PT antimicrobial, antifungal or antitumour compositions, having
 PT increased biological activity
 XX
 PS Claim 30; Page 60; 71pp; English.
 XX
 CC The amphiphilic peptide (SEQ ID NO 61) was prepd. by standard
 CC solid phase methods. The peptide is an analogue of the peptide
 CC having SEQ ID NO 3 (AAR2824), with an N-terminal extension of 8
 CC amino acids. Substitution and deletion analogues of this peptide
 CC have increased biological activity and are effective as pharma-
 CC ceuticals e.g. antibiotics for bacterial, fungal or viral
 CC infections, or in spermicides or antitumour or antiparasitic
 CC agents. Additionally the peptides can be used in wound healing
 CC compsns. or for treating burns or other skin or eye infections.
 CC See also AAR2822-89.
 CC
 SO Sequence 26 AA;
 QY Query Match 58.1%; Score 61; DB 13; Length 26;
 Best Local Similarity 83.3%; Pred. No. 0.25;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 DB 3 KRLKKPPPLKRLKRL 20
 2 KRLKKRLKRLKRLKRL 19
 RESULT 9
 AAR56957
 ID AAR56957 standard; peptide; 15 AA.
 AC AAR56957;
 XX
 DT 17-MAR-1995 (first entry)
 XX
 DE Peptide which neutralises bacterial endotoxin.
 XX
 KM septic shock; bacterial endotoxin; lipopolysaccharide; LPS;
 KM gram negative bacteria; conjugate moiety; septicemia; neutralising;
 KM longer activity; polyvinylpyrrolidone; dextran; beta starch;
 KM polyvinyl alcohol; ion-channel forming; amphiphilic.
 XX
 OS Synthetic.
 XX
 PM WO9413697-A.
 XX
 PD 23-JUN-1994.
 XX
 PF 06-DEC-1993; 93WO-US11841.
 XX
 PR 07-DEC-1992; 92US-0987443.
 XX
 PA (MAGA-) MAGAININ PHARM INC.
 XX
 PI Hendi M, Rao M, Williams TJ;
 DR WPI; 1994-217804/26.
 XX
 PT New conjugates of bioactive amphiphilic peptide(s) and conjugate
 PT moiety - are useful for treatment of septic shock
 XX
 PS Disclosure; Page 120; 141pp; English.
 XX

CC Septic shock is often due to the body's reaction to foreign
 CC lipopolysaccharide (LPS). The compounds of the invention neutralise
 CC bacterial endotoxins without neutralising essential proteins in the
 CC plasma of patients, eg. heparins. They also have longer duration of
 CC activity than unconjugated peptides. In general peptides such as this
 CC are ion-channel forming peptides. The compounds are biologically active
 CC peptides linked to a conjugate moiety, eg. carbohydrates, proteins,
 CC polyvinylpyrrolidone, polyalkylene glycols and polyvinyl alcohols.
 CC The conjugate moiety may be linked at the C- or N-terminal or
 CC internally of the peptide. AAR5591-631 and AAR56879-957 are examples
 CC of these peptide-conjugate moiety compounds
 CC
 SO Sequence 15 AA;
 QY Query Match 53.8%; Score 56.5; DB 15; Length 15;
 Best Local Similarity 83.3%; Pred. No. 0.58;
 Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
 DB 3 KRLKKPPPLKRLKRL 20
 1 KRLKK--LKLKRLKRL 15
 RESULT 10
 ABB74249
 ID ABB74249 standard; Peptide; 16 AA.
 XX
 AC ABB74249;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Amphiphilic fusogenic peptide SEQ ID NO:15.
 XX
 KM Fusogenic; nuclear localisation signal; NLS; encapsulation; lipogene;
 KM liposome; micelle; karyophilic; cytosolic; antitumour; solid tumour;
 KM peptide-lipid-polynucleotide complex; neoplastic disease; gene therapy;
 KM breast carcinoma; prostate carcinoma.
 XX
 OS Synthetic.
 XX
 PM WO200193836-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 08-JUN-2001; 2001WO-US18657.
 XX
 PR 09-JUN-2000; 2000US-210925P.
 XX
 PA (BOUL-) BOULIKAS T.
 XX
 PI Boulikas T;
 DR WPI; 2002-164295/21.
 XX
 PT Encapsulation of plasmid DNA (lipogenes) and therapeutic agents with
 PT nuclear localisation signal/fusogenic peptide conjugates into targeted
 PT liposome complexes -
 XX
 PS Example; Page 38; 107pp; English.
 XX
 CC The present invention describes a method for producing micelles with
 CC entrapped therapeutic agents. The method comprises: (1) combining
 CC negatively charged agent with a cationic lipid in a ratio where 30-90 %
 CC of the negatively charged atoms are neutralised by positive charges on
 CC lipid molecules to form an electrostatic micelle complex in 20-80 %
 CC ethanol; and (2) combining the micelle complex of (a) with fusogenic-
 CC karyophilic peptide conjugates in a 0.0-0.3 ratio, therefore producing
 CC micelles with entrapped therapeutic agents. Also described is a method
 CC for delivering a therapeutic agent in vivo, comprising the administration
 CC of the micelle. ABB74256 to ABB74858 represent specifically claimed
 CC nuclear localisation signal (NLS) peptides for use in the method as the
 CC fusogenic-karyophilic peptides. The micelles produced can have cytostatic
 CC and antitumour activities. The peptide-lipid-polynucleotide complexes

CC produced are useful for inhibiting the progression of neoplastic
 CC diseases. The invention relates to the field of gene therapy and is
 CC directed toward methods for producing peptide-lipid-poly nucleotide
 CC complexes suitable for delivery of polynucleotides. The encapsulated
 CC molecules display therapeutic efficacy in eradicating solid tumours
 CC including but not limited to breast carcinoma or prostate carcinoma.
 CC ABB74235 to ABB74255 are used in the exemplification of the present
 CC invention.

SO Sequence 16 AA;

Query Match 53.8%; Score 56.5; DB 23; Length 16;
 Best Local Similarity 83.3%; Pred. No. 0.61;
 Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 KKLKKPPPLKKLKL 20
 ||||| |||||
 Db 2 KKLKK---LKKLKLKL 16

RESULT 11

AA84924 ID AAR84924 standard; peptide; 40 AA.

AC AAR84924;

DT 18-MAR-1996 (first entry)

DE Alpha-helix-forming oligopeptide (LKKL)10.

KW Alpha-helix; secondary structure; nucleic acid transfer; cationic;
 KM DNA binding peptide; gene therapy; encapsulation.

OS Synthetic.

FT Key Location/Qualifiers
 FT Region 1..4
 FT /label= LKKL
 FT /note= "one of 10 repeat units"

FR2715847-AI.

PD 11-AUG-1995.

PE 08-FEB-1994; 94FR-0001381.

PR 08-FEB-1994; 94FR-0001381.

PA (RHON) RHONE POULENC RORER SA.

PI Bazile D, Emile C, Helene C, Spenlehauer G;

DR WPI; 1995-276981/37.

PT Complex of nucleic acid and oligopeptide with sec. structure - and
 PT transfer vectors contg. them, useful for efficient transfer of
 PT nucleic acid to cells in gene therapy.

PS Claim 6 and Example 1.1; Page 7; 20pp; French.

CC The present peptide is a specific example of a cationic oligopeptide
 CC corresp. to the formula (b-1-1-b)n, where b is a hydrophobic amino acid,
 CC 1 is a hydrophilic amino acid and n is at least 4. In this case, where
 CC b is Leu, 1 is Lys and n = 10, the oligopeptide forms an alpha-helix
 CC which forms a stable complex with a nucleic acid. The complex
 CC is suitable for transferring nucleic acid, esp. in gene therapy.

SO Sequence 40 AA;

Query Match 53.8%; Score 56.5; DB 16; Length 40;
 Best Local Similarity 83.3%; Pred. No. 1.4;
 Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 KKLKKPPPLKKLKL 20
 ||||| |||||
 Db 2 KKLKK---LKKLKLKL 16

RESULT 12

AA66517 ID AAM66517 standard; peptide; 16 AA.

AC AAM66517;

DT 25-NOV-1998 (first entry)

DE Amphiphilic peptide.

KW magainin; analogue; antimicrobial; antitumour; wound healing;
 KM Cpf; amphiphilic; xfp peptide.

OS Synthetic.

FT Key Location/Qualifiers
 FT Misc-difference 1..16
 FT /note= "D-form residues"

PN US5792831-A.

PD 11-AUG-1998.

PE 17-NOV-1994; 94US-0343882.

PR 05-OCT-1993; 93US-0133740.

PR 08-FEB-1990; 90US-0476629.

PR 14-MAY-1990; 90US-0522688.

PR 28-APR-1992; 92US-0874685.

PR 17-NOV-1994; 94US-0343882.

PA (MAGA-) MAGANIN PHARM INC.

PI Maloy WL;

DR WPI; 1998-456190/39.

PT Magainin peptide analogues - useful as antimicrobial or antitumour
 PT agents, etc.

PS Disclosure; Column 18; 25pp; English.

CC The invention relates to analogues of a magainin I or II, D-form
 CC analogues, deletion analogues or related peptides. It also relates
 CC to basic polypeptides having at least 16 amino acids, including at least
 CC 8 hydrophobic amino acids and at least 8 hydrophilic amino acids. The
 CC peptides may be used as antimicrobial agents, antiviral agents,
 CC antibiotics, antitumour agents, antiparasitic agents, spermicides,
 CC preservatives or sterilants, or agents for promoting wound healing. The
 CC present sequence represents a specific example of a peptide disclosed in
 CC the specification.

SO Sequence 16 AA;

Query Match 52.4%; Score 55; DB 19; Length 16;
 Best Local Similarity 75.0%; Pred. No. 0.96;
 Matches 15; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

OY 1 KKKLKKPPPLKKLKL 20
 ||||| |||||
 Db 1 KKLKKLKK---LKKLKLKL 16

RESULT 13

AA21385 ID AAR21385 standard; peptide; 17 AA.

AC AAR21385;

XX 16-MAY-1992 (first entry)
 DT XX
 DE Sequence of amphiphilic peptide SEQ ID No. 48 with
 DE C-terminal amide and acetylated at N-terminus.
 XX
 KM Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;
 KM spermicide; wound healing; steriliant.
 XX
 PN WO9201462-A.
 XX
 PD 06-FEB-1992.
 XX
 PF 17-JUL-1991; 91WO-US05047.
 XX
 PR 19-JUL-1990; 90US-0554422.
 XX
 PA (SCRI-) SCRIPPS RES INST.
 XX
 PI Houghten RA, Blondelle S;
 XX
 DR WPI: 1992-064700/08.
 XX
 PT Method for inhibiting target cell and virus growth - comprises
 PT administering amphiphilic peptide compns, useful for treating
 PT viral and phytopathogenic infections, tumours and burns
 XX
 PS Claim 4; Page 54; 72pp; English.
 XX
 CC The peptides of the invention are effective pharmaceuticals having
 CC anti-microbial, anti-viral and anti-tumour activity. They are also
 CC useful for inhibiting, preventing or destroying the motility of
 CC sperm and hence have application in a spermicide preparation. They
 CC also have anti-parasitic activity and are useful in wound healing,
 CC as preservatives and sterilants and to inhibit growth of
 CC phytopathogenic microorganisms. AAR20969 and AAR20970 were published in
 CC Highten and Ostresh, Bio Chromatography, Vol 2, Issue 2, page 80-83,
 CC 1987.
 CC
 XX
 SQ Sequence 17 AA:
 QY
 DB 1 KKKKLLKKPPPLKKLLKKL 20
 2 KKKKLLKK---LKKLLKKL 17

Query Match 52.4%; Score 55; DB 13; Length 17;
 Best Local Similarity 75.0%; Pred. No. 1;
 Matches 15; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

RESULT 14
 AAR22869
 ID AAR22869 standard; Peptide: 17 AA.
 XX
 AC AAR22869;
 XX
 DT 22-AUG-1992 (first entry)
 XX
 DE Amphiphilic peptide to inhibit growth of a target cell.
 XX
 KM Antimicrobial; antifungal; antitumour; pharmaceuticals; antibiotics;
 KM antiparasitic; spermicides; burns; wound healing.
 XX
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "acylated"
 FT Modified-site 17 /note= "amidated"
 FT
 XX CA2047317-A.
 XX
 XX

PD 20-JAN-1992.
 XX
 PF 18-JUL-1991; 91CA-2024317.
 XX
 PR 08-JUL-1991; 91US-0725331.
 PR 19-JUL-1990; 90US-0554422.
 XX
 PA (SCRI-) SCRIPPS RES INST.
 XX
 PI Houghten RA, Blondelle S;
 XX
 DR WPI: 1992-114943/15.
 XX
 PF Amphiphilic peptide(s) and analogues - for use in e.g.
 PT antimicrobial, antifungal and antitumour compositions, having
 PT increased biological activity
 XX
 PS Claim 9; Page 54; 71pp; English.
 XX
 CC The amphiphilic peptide (SEQ ID NO 48) was prepd. by standard
 CC solid phase methods. The peptide is an analogue of the peptide
 CC having SEQ ID NO 3 (AAR22824), with Leu at position 4 deleted.
 CC Substitution and deletion analogues of this peptide have
 CC increased biological activity and are effective as pharmaceuticals
 CC e.g. antibiotics for bacterial, fungal or viral infections, or in
 CC spermicides or antitumour or antiparasitic agents. Additionally
 CC the peptides can be used in wound healing compns. or for treating
 CC burns or other skin or eye infections.
 CC See also AAR22822-89.
 CC
 XX
 SQ Sequence 17 AA:
 QY
 DB 1 KKKKLLKKPPPLKKLLKKL 20
 2 KKKKLLKK---LKKLLKKL 17

Query Match 52.4%; Score 55; DB 13; Length 17;
 Best Local Similarity 75.0%; Pred. No. 1;
 Matches 15; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

RESULT 15
 AAR21360
 ID AAR21360 standard; peptide: 18 AA.
 XX
 AC AAR21360;
 XX
 DT 16-MAY-1992 (first entry)
 XX
 DE Sequence of amphiphilic peptide SEQ ID No. 23 which may be a C-
 DE terminal amide and/or may be acetylated at N-terminus.
 XX
 KM Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;
 KM spermicide; wound healing; steriliant.
 XX
 PN WO9201462-A.
 XX
 PD 06-FEB-1992.
 XX
 PF 17-JUL-1991; 91WO-US05047.
 XX
 PR 19-JUL-1990; 90US-0554422.
 XX
 PA (SCRI-) SCRIPPS RES INST.
 XX
 PI Houghten RA, Blondelle S;
 XX
 DR WPI: 1992-064700/08.
 XX
 PT Method for inhibiting target cell and virus growth - comprises
 PT administering amphiphilic peptide compns, useful for treating
 PT viral and phytopathogenic infections, tumours and burns
 XX

PS Claim 4; Page 42; 72pp; English.

XX
CC The peptides of the invention are effective pharmaceuticals having
CC anti-microbial, anti-viral and anti-tumour activity. They are also
CC useful for inhibiting, preventing or destroying the motility of
CC sperm and hence have application in a spermicide preparation. They
CC also have anti-parasitic activity and are useful in wound healing,
CC as preservatives and sterilants and to inhibit growth of
CC phytopathogenic microorganisms. AAR20969 and AAR20970 were published in
CC Haighten and Ostresh, Bio Chromatography, Vol 2, Issue 2, page 80-83,
CC 1987.

XX
SQ Sequence 18 AA;

Query Match 52.4%; Score 55; DB 13; Length 18;

Best Local Similarity 77.8%; Pred. No. 1.1;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 KKLKKPPPLKKLKL 20
||| |
Db 1 KKLKKLKKLKKLKL 18

Search completed: June 6, 2003, 13:23:39
Job time : 31.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 13:22:11 ; Search time 12.5 Seconds
(without alignments)
153.815 Million cell updates/sec

Title: us-10-081-418-2
Perfect score: 105
Sequence: 1 KKKLLKKPPPLKKLLKL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	48.6	369	T24797	hypothetical prote
2	50	47.6	900	T47732	probable translati
3	50	47.6	900	T50773	translacion initia
4	49	46.7	110	D72863	Acort-107 protein
5	49	46.7	340	E81331	uroporphyrinogen d
6	49	46.7	656	S69554	hypothetical prote
7	49	46.7	1235	S24109	phosphorylase kina
8	48.5	46.2	668	C24785	hypothetical prote
9	48.5	46.2	700	C1H0H2	calpain (BC 3.4.22
10	48	45.7	70	S70028	mitogen-activated
11	48	45.7	339	E71918	uroporphyrinogen d
12	48	45.7	340	D64595	uroporphyrinogen d
13	48	45.7	867	A81707	ATP-dependent Clp
14	48	45.7	2897	B48666	cell proliferation
15	48	45.7	3256	A48666	cell proliferation
16	47	44.8	1324	S06187	RNA2 polypeptid
17	46	43.8	85	D69386	conserved hypotet
18	46	43.8	157	S49448	oleosin - rape
19	46	43.8	537	F0MWV7	gag polyprotein -
20	46	43.8	1341	S66835	probable membrane
21	45.5	43.3	355	B75195	hypothetical prote
22	45.5	43.3	479	A64117	serine-type D-Ala-
23	45.5	43.3	700	S38361	calpain (BC 3.4.22
24	45	42.9	151	S07883	hypothetical prote
25	45	42.9	165	S50195	oleosin - rape
26	45	42.9	375	S58484	gag protein - maiz
27	45	42.9	493	A85433	sugar transporter
28	45	42.9	563	T40705	conserved hypotet
29	45	42.9	1152	H88533	probable ATP-depen

30	44.5	42.4	629	2	A71023	arginine-tRNA ligase
31	44	41.9	91	2	T08131	oleosin-like prote
32	44	41.9	143	2	JC4203	nitrogenase NifX c
33	44	41.9	143	2	T09232	nifX protein - Fra
34	44	41.9	186	2	T45838	hypothetical prote
35	44	41.9	224	2	F96668	protein FIN19.5 (1
36	44	41.9	319	2	S34806	hypothetical prote
37	44	41.9	469	2	JC5798	F0P1-ATPase (EC 3.
38	44	41.9	592	2	JC4642	purh bifunctional
39	44	41.9	681	2	C84612	hypothetical prote
40	44	41.9	779	2	S56245	cell division cont
41	44	41.9	781	2	JC7382	DNA-directed DNA p
42	44	41.9	879	2	A56277	DNA-directed DNA p
43	44	41.9	893	2	F72253	hypothetical prote
44	44	41.9	1073	1	T08328	plasmid replicatio
45	44	41.9	1309	1	S35484	peptidyl-dipectida

ALIGNMENTS

RESULT 1
T24797
hypothetical protein T10C6.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T24797
R:McMurray, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: 219936
A:Accession: T24797
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-369 <WILD>
A:Cross-references: EMBL:Z93388; PIDN:GAB07663.1; GSPDB:GN00023; CESP:T10C6.7
A:Experimental source: clone T10C6
C:Genetics:
A:Gene: CESP:T10C6.7
A:Map position: 5
A:Introns: 44/2; 114/2; 219/3; 313/2
C:Superfamily: Caenorhabditis elegans hypothetical protein F08F3.6

Query Match 48.6% Score 51; DB 2; Length 369;
Best Local Similarity 50.0% Pred. No. 11;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKLLKKPPPLKKLLKL 20
DB 20 KKKRLQLPPPAFANILKIL 39

RESULT 2
T47732
probable translation initiation factor eIF-3 chain 8 F18021.110 [Imported] - Arabidop
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 05-May-2000
C:Accession: T47732
R:Benes, V.; Wurmach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Rudd, S.; Lemcke, J.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24474
A:Accession: T47732
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-900 <BEN>
A:Cross-references: EMBL:ATP18021
A:Experimental source: cultivar Columbia; BAC clone F18021
C:Genetics:
A:Map position: 3
A:Introns: 8/3; 593/3
A:Note: F18021.110

Query Match 47.6% Score 50; DB 2; Length 900;
Best Local Similarity 50.0% Pred. No. 37;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
OY 2 WKLLKKRPPPLKKLKK 19
DB 215 WEKMLSKDKLLEKLMNK 232

RESULT 3

T50773
translation initiation factor homolog p105 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50773
R:Karniol, B.; Yabaiom, A.; Krok, S.; Tsuge, T.; Matsui, M.; Deng, X.W.; Chamovitz, D.A.
FEBS Lett. 439, 173-179, 1998
A:Title: The Arabidopsis homolog of an eIF3 complex subunit associates with the COP9
A:Reference number: Z24448; MUID:9849901; PMID:9849901
A:Accession: T50773
A:Status: Preliminary;
A:Molecule type: mRNA
A:Residues: 1-900 <RAB>
A:Cross-references: EMBL:AF040102; PIDN:AAC83464.1

Query Match 47.6%; Score 49; DB 2; Length 900;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 WKLLKKRPPPLKKLKK 19

DB 215 WEKMLSKDKLLEKLMNK 232

RESULT 4

D72863
AcOrf-107 protein - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, ACMPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C:Accession: D72863
R:AYres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Accession: D72863
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-110 <AVR>
A:Cross-references: GB:L22858; NID:9510708; PIDN:AAA66737.1; PID:9559176
C:Genetics:
A:Gene: AcOrf-107

Query Match 46.7%; Score 49; DB 2; Length 110;
Best Local Similarity 64.7%; Pred. No. 5.7;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 KKLKKRPPPLKKLKK 19

DB 85 KCLLPAPPPOLKLEKK 101

RESULT 5

E81331
uroporphyrinogen decarboxylase (EC 4.1.1.37) Cj1243 [imported] - Campylobacter jejuni (S
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: E81331
R:parhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churche, C.; Basham, D.; Chitt
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h
A:Reference number: AB1250; MUID:20150912; PMID:10688204
A:Accession: E81331
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-340 <PAR>
A:Cross-references: GB:AL139077; GB:AL111168; NID:96968444; PIDN:CAB73497.1; PID:9696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: heme; Cj1243
C:Superfamily: uroporphyrinogen decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 46.7%; Score 49; DB 2; Length 340;
Best Local Similarity 55.0%; Pred. No. 19;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKLKKRPPPLKKLKK 20

DB 160 KCKKMLYQNPDLKILKK 179

RESULT 6

S69554
hypothetical protein YDR496c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 19-Apr-2002
C:Accession: S69554
R:Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A:Description: The sequence of S. cerevisiae cosmid 8166, 9787, 9717, and lambda 307
A:Reference number: S69553
A:Accession: S69554
A:Molecule type: DNA
A:Residues: 1-656 <DIE>
A:Cross-references: EMBL:U33057; NID:9927764; PID:9927766; GSPDB:GN00004; MIPS:YDR496
A:Gene: SGD:PUF6; MIPS:YDR496c
A:Cross-references: S69554
A:Map position: 4R

Query Match 46.7%; Score 49; DB 2; Length 656;
Best Local Similarity 47.4%; Pred. No. 37;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 2 WKLLKKRPPPLKKLKK 20

DB 131 WERLKVKTPLPKQIREKL 149

RESULT 7

S24109
phosphorylase kinase (EC 2.7.1.38) alpha chain, hepatic - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S24109
R:Davidson, J.J.; Oezcelik, T.; Hamacher, C.; Williams, P.J.; Francke, U.; Killmann, M
Proc. Natl. Acad. Sci. U.S.A. 89, 2096-2100, 1992
A:Title: cDNA cloning of a liver isoform of the phosphorylase kinase alpha subunit an
A:Reference number: S24109; MUID:92196064; PMID:1372435
A:Accession: S24109
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1235 <DAV>
A:Cross-references: EMBL:X60421; NID:91657; PIDN:CMA42952.1; PID:91658
C:Keywords: liver; phosphotransferase

Query Match 46.7%; Score 49; DB 2; Length 1235;
Best Local Similarity 56.2%; Pred. No. 71;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 LKKRPPPLKKLKK 20

DB 734 LVDSFQPLKKLKK 749

RESULT 8

C24785
 Hypothetical protein 668 - slime mold (Dictyostelium discoideum) transposon DIRS-1 (frag
 C:Species: Dictyostelium discoideum
 C>Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 18-Jun-1993
 C:Accession: C24785
 R:Capello, J.; Handelsman, K.; Lodish, H.F.
 A>Title: Sequence of Dictyostelium DIRS-1: an apparent retrotransposon with inverted ter
 A:Reference number: A94654; MUID:86079481; PMID:2416457
 A:Accession: C24785
 A:Molecule type: DNA
 A:Residues: 1-668 <CAP>

Query Match 46.2%; Score 48.5; DB 2; Length 668;
 Best Local Similarity 47.6%; Pred. No. 44;
 Matches 10; Conservative 5; Mismatches 5; Indels 1; Gaps 1;
 Oy 1 KKKLKKPP-LLKKLKL 20
 Db 572 QWQCLAFPPILPILSLK 592

RESULT 9

C1H0H2
 calpain (EC 3.4.22.17) large chain 2 - human
 N:Alternate names: calpain chain L-2; calpain II catalytic chain; high-calcium requiring
 C:Species: Homo sapiens (man)
 C>Date: 21-Nov-1993 #sequence_revision 09-Aug-1997 #text_change 16-Jul-1999
 C:Accession: S10590; A31218; A35529
 R:Sortimachi, H.; Ohmi, S.; Emori, Y.; Kawasaki, H.; Saito, T.C.; Ohno, S.; Minami, Y.; S
 Biol. Chem. Hoppe-Seyler 371(Suppl.), 171-176, 1990
 A>Title: A novel member of the calcium-dependent cysteine protease family.
 A:Reference number: S10589; MUID:90380278; PMID:2400579
 A:Accession: S10590
 A:Molecule type: mRNA
 A:Residues: 1-700 <SOR>
 R:Imajo, S.; Aoki, K.; Ohno, S.; Emori, Y.; Kawasaki, H.; Sugihara, H.; Suzuki, K.
 Biochemistry 27, 8122-8128, 1988
 A>Title: Molecular cloning of the cDNA for the large subunit of the high-Ca(2+)-regulin
 A:Reference number: A31218; MUID:89166474; PMID:2852952
 A:Accession: A31218
 A:Molecule type: mRNA; protein
 A:Residues: 1-210, 'I', 212-394, 'D', 396-445, 'I', 447-700 <TMA>
 A:Cross-references: GB:M23254; NID:9511636; PIDN:AAA55645.1; PID:9511637
 A:Note: parts of this sequence were determined by protein sequencing; the amino end of t
 J. Biol. Chem. 264, 6404-6411, 1989
 A>Title: Tandemly reiterated negative enhancer-like elements regulate transcription of a
 A:Reference number: A35529; MUID:89197947; PMID:2539381
 A:Accession: A35529
 A:Molecule type: DNA
 A:Residues: 1-67, 'G', 69-72, 'T', 75-78, 'R' <HAT>
 A:Cross-references: DDBJ:J04700; NID:9179910; PIDN:AAA52760.1; PID:9463086
 C:Genetics: GDB:CAPN2; mCAP; CAPM1
 A:Gene: GDB:CAPN2; mCAP; CAPM1
 A:Cross-references: GDB:119750; OMIM:114230
 A:Map position: 1pter-1qter
 C:Complex: heterodimer of L (large) and S (small) chains
 C:Function:
 A:Description: catalyzes the hydrolysis of peptides
 A:Note: Cleaves preferentially after tyrosine, methionine, or arginine residues and bef
 C:Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C:Keywords: acetylated amino end; calcium binding; cysteine proteinase; EF hand; heterod
 F:2-700/Product: calpain large chain 2 #status predicted <CAP>
 F:75-327/Domain: calpain catalytic domain homology <CAP>
 F:529-560/Domain: calmodulin repeat homology <EF1>
 F:572-604/Domain: calmodulin repeat homology <EF2>
 F:605-634/Domain: calmodulin repeat homology <EF3>
 F:637-668/Domain: calmodulin repeat homology <EF4>
 F:2-modified site: acetylated amino end (Ala) (in mature form) #status predicted
 F:105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 46.2%; Score 48.5; DB 1; Length 700;

Best Local Similarity 52.6%; Pred. No. 46;
 Matches 10; Conservative 5; Mismatches 3; Indels 1; Gaps 1;
 Oy 1 KKKLKKPP-LLKKLKL 19
 Db 213 EWE-LKKPPNLEKTIK 230

RESULT 10

S70028
 mitogen-activated protein kinase kinase 1 - common carp (fragments)
 C:Species: Cyprinus carpio (common carp)
 C>Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 17-Mar-1999
 C:Accession: S70028
 R:Leu, J.H.; Lee, M.S.; Chen, K.T.; Chang, G.D.; Chou, C.K.; Huang, C.J.
 Biochim. Biophys. Acta 1306, 133-136, 1996
 A>Title: Genomic structure of carp mitogen-activated protein kinase kinase 1 gene.
 A:Reference number: S70028; MUID:96221282; PMID:8634328
 A:Accession: S70028
 A:Status: Preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-7;8-14;15-21;22-28;29-35;36-42;43-49;50-56;57-63;64-70 <LEU>

Query Match 45.7%; Score 48; DB 2; Length 70;
 Best Local Similarity 43.8%; Pred. No. 4.9;
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 Oy 5 LKKPPPLKLLKLKL 20
 Db 50 IYNEPPVVKCLMKKM 65

RESULT 11

E71918
 uroporphyrinogen decarboxylase - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Jun-1999
 C:Accession: E71918
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, O.; Taylor, D.E.; Vovis, G.F
 Nature 397, 176-180, 1999
 A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric p
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: E71918
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-339 <ARN>
 A:Cross-references: GB:AE001487; GB:AE001439; NID:94155086; PIDN:AD06123.1; PID:9415
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: hemE
 C:Superfamily: uroporphyrinogen decarboxylase

Query Match 45.7%; Score 48; DB 2; Length 339;
 Best Local Similarity 55.0%; Pred. No. 26;
 Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 Oy 1 KKKLKKPP-LLKKLKL 20
 Db 159 KSKKMLSEPEVKKALDEL 178

RESULT 12

D64595
 uroporphyrinogen decarboxylase - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 22-Jun-1999
 C:Accession: D64595
 R:Tombs, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khakhria, H.G.; Glodek, A.; McK
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wetthey,
 Nature 386, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
 A:Reference number: A64520; MID:97394467; PMID:9252185
 A:Accession: D64595
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-340 <TOM>
 A:Cross-references: GB:AE000574; GB:AE000511; NID:92313721; PIDN:AD07669.1; PID:9231372
 C:Superfamily: uroporphyrinogen decarboxylase

Query Match 45.7%; Score 48; DB 2; Length 340;
 Best Local Similarity 55.0%; Pred. No. 26;
 Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKLKKPPPLKKLKL 20
 |||:| | :|||
 Db 160 KSKKMLYSEPVKALKLEKL 179

RESULT 13

A81707

ATP-dependent Clp proteinase, chain B JTC0389 [Imported] - *Chlamydia muridarum* (strain N1)
 C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* Mopn
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001
 C:Accession: A81707
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gail, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of *Chlamydia trachomatis* Mopn and *Chlamydia pneumoniae* AR39.
 A:Reference number: A81500; MID:20150255; PMID:10684935
 A:Accession: A81707
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-867 <TEU>
 A:Cross-references: GB:AE002306; GB:AE002160; NID:97190428; PIDN:AAF39246.1; PID:9719043

A:Experimental source: strain N1gg (Mopn)
 C:Genetics:
 A:Gene: JTC0389
 C:Superfamily: endopeptidase Clp ATP-binding chain
 C:Keywords: ATP; molecular chaperone; nucleotide binding; P-loop
 F:205-212/Region: nucleotide-binding motif A (P-loop)
 F:273-278/Region: nucleotide-binding motif B
 F:605-612/Region: nucleotide-binding motif A (P-loop)
 F:673-678/Region: nucleotide-binding motif B
 F:211/Binding site: ATP (Lys) #status predicted
 F:611/Binding site: ATP (Lys) #status predicted

Query Match 45.7%; Score 48; DB 2; Length 867;
 Best Local Similarity 47.4%; Pred. No. 66;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 KKKLKKPPPLKKLKL 20
 ||:| | | :||:|
 Db 124 WKRTVKTSPDLKELIKL 142

RESULT 14

B48666

cell proliferation antigen Ki-67, short form - human
 C:Species: *Homo sapiens* (man)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 07-May-1999
 C:Accession: B48666
 R:Schlueter, C.; Duchrow, M.; Mohlenberg, C.; Becker, M.H.G.; Key, G.; Flad, H.D.; Gerde,
 J. Cell Biol. 123, 513-522, 1993
 A:Title: The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiq
 uous.

A:Reference number: A48666; MID:94043435; PMID:8227122
 A:Accession: B48666

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2897 <SCH>
 A:Cross-references: EMBL:X65551
 C:Superfamily: Kinase interaction domain homology

C:Keywords: alternative splicing; cell cycle control; nucleus; tandem repeat
 F:29-91/Domain: kinase interaction domain homology <KIH>

Query Match 45.7%; Score 48; DB 2; Length 2897;
 Best Local Similarity 47.4%; Pred. No. 2,4e+02;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 KKKLKKPPPLKKLKK 19
 ||:| | | :||:|
 Db 176 KRKSLVMTPTVPLKTIKE 194

RESULT 15

A48666

cell proliferation antigen Ki-67, long form - human
 C:Species: *Homo sapiens* (man)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 08-Oct-1999
 C:Accession: A48666
 R:Schlueter, C.; Duchrow, M.; Mohlenberg, C.; Becker, M.H.G.; Key, G.; Flad, H.D.; Ge
 J. Cell Biol. 123, 513-522, 1993

A:Title: The cell proliferation-associated antigen of antibody Ki-67: a very large, u
 biquous.
 A:Reference number: A48666; MID:94043435; PMID:8227122
 A:Accession: A48666
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-3256 <SCH>
 A:Cross-references: EMBL:X65550; NID:9415818; PIDN:CAA46519.1; PID:9415819
 C:Superfamily: kinase interaction domain homology
 C:Keywords: alternative splicing; cell cycle control; nucleus; tandem repeat
 F:29-91/Domain: kinase interaction domain homology <KIH>

Query Match 45.7%; Score 48; DB 2; Length 3256;
 Best Local Similarity 47.4%; Pred. No. 2,7e+02;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 KKKLKKPPPLKKLKK 19
 ||:| | | :||:|
 Db 535 KRKSLVMTPTVPLKTIKE 553

Search completed: June 6, 2003, 13:25:28
 Job time: 14.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein. - protein search, using sw model

Run on: June 6, 2003, 13:22:41 ; Search time 15.5 Seconds

(without alignments)
133.213 Million cell updates/sec

Title: US-10-081-418-2

Perfect score: 105

Sequence: 1 KMKLLKKPPLKKLLKLL 20

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 392085 segs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	20	9	US-10-081-418-2
2	56.5	53.8	16	9	US-09-876-904A-15
3	56.5	53.8	16	9	US-09-876-904A-629
4	53	50.5	17	9	US-09-999-745-55
5	53	50.5	17	9	US-09-554-000-39
6	50	47.6	18	10	US-09-447-966-1
7	49	46.7	26	10	US-09-030-619-125
8	48.5	46.2	23	9	US-09-820-053A-96
9	48.5	46.2	23	9	US-09-820-053A-96
10	48.5	46.2	700	9	US-10-038-557A-9
11	48.5	46.2	700	10	US-08-840-707A-9
12	48.5	46.2	700	10	US-09-768-877-23
13	48.5	46.2	513	9	US-10-051-902-20
14	48.5	46.2	513	9	US-10-051-909-20
15	48	45.7	3256	10	US-09-919-172-98
16	46	43.8	46	12	US-10-036-869-11
17	45	42.9	47	10	US-09-864-761-42909
18	45	42.9	49	10	US-09-864-761-40695
19	45	42.9	54	10	US-09-864-761-48671

20	45	42.9	74	10	US-09-864-761-33838	Sequence 33838, A
21	45	42.9	510	9	US-10-051-902-22	Sequence 22, Appl
22	45	42.9	510	9	US-10-051-909-22	Sequence 22, Appl
23	44.5	42.4	20	9	US-09-820-053A-68	Sequence 68, Appl
24	44.5	42.4	610	9	US-09-820-053A-68	Sequence 4495, Ap
25	44	41.9	22	9	US-09-820-053A-46	Sequence 46, Appl
26	44	41.9	22	9	US-09-820-053A-46	Sequence 48, Appl
27	44	41.9	22	9	US-09-820-053A-46	Sequence 55, Appl
28	44	41.9	22	9	US-09-820-053A-46	Sequence 17, Appl
29	44	41.9	67	10	US-09-843-845-17	Sequence 609, App
30	44	41.9	232	10	US-10-051-902-28	Sequence 28, Appl
31	44	41.9	529	9	US-10-051-902-28	Sequence 28, Appl
32	44	41.9	529	9	US-10-051-902-28	Sequence 1218, Ap
33	44	41.9	598	10	US-09-925-301-1218	Sequence 28, Appl
34	44	41.9	907	9	US-10-060-019-29	Sequence 29, Appl
35	44	41.9	907	9	US-10-200-154-4	Sequence 4, Appl1
36	44	41.9	940	9	US-09-954-043-4	Sequence 4, Appl1
37	43	41.0	23	9	US-10-102-806-678	Sequence 678, App
38	43	41.0	23	9	US-09-820-053A-9	Sequence 9, Appl
39	43	41.0	26	9	US-09-820-053A-161	Sequence 161, App
40	43	41.0	26	10	US-09-808-139-2	Sequence 122, App
41	43	41.0	216	12	US-10-057-510-4	Sequence 4, Appl1
42	43	41.0	242	10	US-09-764-898-199	Sequence 199, App
43	43	41.0	489	10	US-09-956-425-6	Sequence 6, Appl1
44	43	41.0	491	9	US-09-029-327-2	Sequence 2, Appl1
45	43	41.0	491	10	US-09-886-077-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-081-418-2
Sequence 2, Application US/10081418
Publication No. US20030096745A1
GENERAL INFORMATION:
APPLICANT: HAHM, Kyung-Soo
APPLICANT: LEE, Dong Gun
APPLICANT: PARK, Yoonkyung
APPLICANT: KIM, Hee Nam
TITLE OF INVENTION: NO. US20030096745A1 peptides with increased + charge and hydrophobicity by substituting one or more amino acids of CA-M peptide and
TITLE OF INVENTION: Pharmaceutical compositions containing thereof
FILE REFERENCE: 428.1014
CURRENT APPLICATION NUMBER: US/10/081,418
CURRENT FILING DATE: 2002-02-22
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Kopatentia 1.71
SEQ ID NO 2
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide with increased + charge and hydrophobicity by substituting amino acids of SEQ. ID. NO 1 with lysine and leucine
US-10-081-418-2

Query Match
Best Local Similarity 100.0%; Score 105; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMKLLKKPPLKKLLKLL 20
DB 1 KMKLLKKPPLKKLLKLL 20

RESULT 2
US-09-876-904A-15
Sequence 15, Application US/09876904A
Publication No. US20030072794A1
GENERAL INFORMATION:
APPLICANT: BOULIKAS, TENI
TITLE OF INVENTION: ENCAPSULATION OF PLASMID DNA (LIPOGENES TM) AND THERAPEUTIC

TITLE OF INVENTION: AGENTS WITH NUCLEAR LOCALIZATION SIGNAL/FUSOGENIC PEPTIDE
FILE REFERENCE: TB-2002.00
CURRENT APPLICATION NUMBER: US/09/876,904A
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 60/210,925
PRIOR FILING DATE: 2000-06-09
NUMBER OF SEQ ID NOS: 629
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic amphiphilic
US-09-876-904A-15

Query Match 53.8%; Score 56.5; DB 9; Length 16;
Best Local Similarity 83.3%; Pred. No. 0.25;
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 KKLKKPPPLKKLKL 20
DB 2 KKLKK---LKKLKLKL 16

RESULT 3
US-09-876-904A-629
Sequence 629, Application US/09876904A
Publication No. US20030072794A1
GENERAL INFORMATION:
APPLICANT: BOULIKAS, TENI
TITLE OF INVENTION: ENCAPSULATION OF PLASMID DNA (LIPIDGENES TM) AND THERAPEUTIC
FILE REFERENCE: TB-2002.00
CURRENT APPLICATION NUMBER: US/09/876,904A
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 60/210,925
PRIOR FILING DATE: 2000-06-09
NUMBER OF SEQ ID NOS: 629
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 629
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
US-09-876-904A-629

Query Match 53.8%; Score 56.5; DB 9; Length 16;
Best Local Similarity 83.3%; Pred. No. 0.25;
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 KKLKKPPPLKKLKL 20
DB 2 KKLKK---LKKLKLKL 16

RESULT 4
US-09-999-745-55
Sequence 55, Application US/09999745
Patent No. US20020157120A1
GENERAL INFORMATION:
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
APPLICANT: Tsien, Roger Y.
APPLICANT: Balid, Geoffrey
TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
FILE REFERENCE: REGEN1470-1
CURRENT APPLICATION NUMBER: US/09/999,745
CURRENT FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: 09/316,920

PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.0
SEQ ID NO 55
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Calmodulin binding peptide-2
US-09-999-745-55

Query Match 50.5%; Score 53; DB 9; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.75;
Matches 13; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

OY 2 WKLLKKPPPLKKLKL 18
DB 3 WKLLKK---LKKLKLKL 15

RESULT 5
US-09-554-000-39
Sequence 39, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000
CURRENT FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 08/818,252
PRIOR FILING DATE: 1997-03-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Calmodulin binding peptide-2
US-09-554-000-39

Query Match 50.5%; Score 53; DB 9; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.75;
Matches 13; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

OY 2 WKLLKKPPPLKKLKL 18
DB 3 WKLLKK---LKKLKLKL 15

RESULT 6
US-09-447-966-1
Sequence 1, Application US/09447966
Patent No. US20010004636A1
GENERAL INFORMATION:
APPLICANT: WOLFF, JON A
APPLICANT: Monahan, Sean D
APPLICANT: Slatum, Paul M E
APPLICANT: Hagstrom, James E
APPLICANT: Budker, Vladimir G
APPLICANT: Rozena, David B
TITLE OF INVENTION: INTRAVASCULAR DELIVERY OF NON-VIRAL NUCLEIC ACID
FILE REFERENCE: Mirus.013.03
CURRENT APPLICATION NUMBER: US/09/447,966
CURRENT FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: 60/121730
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: 60/146564
PRIOR FILING DATE: 1999-07-30
NUMBER OF SEQ ID NOS: 1

```

; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 18-mer
; OTHER INFORMATION: positive charge
US-09-447-966-1

Query Match
Best Local Similarity 47.6%; Score 50; DB 10; Length 18;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 KKKKKPPPLKKLKL 20
    |||| | |||||
Db 1 KKKKKLKKLKKLKKL 17

RESULT 7
US-09-030-619-125
; Sequence 125, Application US/090306195
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 125
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-125

Query Match
Best Local Similarity 46.7%; Score 49; DB 10; Length 26;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 KKKKKLKKPPPLKKLKL 17
    || | | | | | | |
Db 1 KKKLPLKKLTPAVKKVL 17

RESULT 8
US-09-820-053A-49
; Sequence 49, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:
; APPLICANT: Owen, Donald R.
; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
; FILE REFERENCE: HELX027
; CURRENT APPLICATION NUMBER: US/09/820,053A
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 49
; LENGTH: 23
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC SEQUENCE
```

```

; NAME/KEY: MOD_RES
; LOCATION: (23)
; OTHER INFORMATION: AMIDATION
US-09-820-053A-49

Query Match
Best Local Similarity 46.2%; Score 48.5; DB 9; Length 23;
Matches 12; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

OY 1 KKKKKLKKPPPLKKLKL 20
    || || | | | | |
Db 1 KW-KLKKKTKLKKFPAKL 19

RESULT 9
US-09-820-053A-56
; Sequence 56, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:
; APPLICANT: Owen, Donald R.
; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
; FILE REFERENCE: HELX027
; CURRENT APPLICATION NUMBER: US/09/820,053A
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 56
; LENGTH: 23
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC SEQUENCE
US-09-820-053A-56

Query Match
Best Local Similarity 46.2%; Score 48.5; DB 9; Length 23;
Matches 12; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

OY 1 KKKKKLKKPPPLKKLKL 20
    || || | | | | |
Db 1 KW-KLKKKTKLKKFPAKL 19

RESULT 10
US-10-038-557A-9
; Sequence 9, Application US/10038557A
; Publication No. US20030092664A1
; GENERAL INFORMATION:
; APPLICANT: Fredelking, Terry M.
; APPLICANT: Ignatyev, George M.
; TITLE OF INVENTION: INFECTIONS AND OTHER DISORDERS
; FILE REFERENCE: 24861-301D
; CURRENT APPLICATION NUMBER: US/10/038,557A
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 09/840,707
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/562,979
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/198,210
; PRIOR FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: IL-1 receptor intracellular ligand protein
; OTHER INFORMATION: comprising amino acid sequence
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 5,817,476
; PATENT FILING DATE: 1995-06-07
```

PUBLICATION DATE: 1998-10-06
US-10-038-557A-9

Query Match 46.2%; Score 48.5; DB 9; Length 700;
Best Local Similarity 52.6%; Pred. No. 1.2e+02;
Matches 10; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

OY 1 KKKLLKKPPPLKKLLK 19
DB 213 EWE-LKKPPNLFKIIK 230

RESULT 11
US-09-840-707A-9
Sequence 9, Application US/09840707A
Patent No. US2002007276A1
GENERAL INFORMATION:

APPLICANT: Fredeking, Terry M.
APPLICANT: Ignatyev, George M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
FILE REFERENCE: 24881-301C
CURRENT APPLICATION NUMBER: US/09/840,707A

PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/562,979
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/198,210
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
OTHER INFORMATION: comprising amino acid sequence
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,817,476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-09-840-707A-9

Query Match 46.2%; Score 48.5; DB 10; Length 700;
Best Local Similarity 52.6%; Pred. No. 1.2e+02;
Matches 10; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

OY 1 KKKLLKKPPPLKKLLK 19
DB 213 EWE-LKKPPNLFKIIK 230

RESULT 12
US-09-768-877-23

Sequence 23, Application US/09768877
Patent No. US20020150896A1
GENERAL INFORMATION:

APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEANUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.

TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/768,877
CURRENT FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 09/422,869
PRIOR FILING DATE: 1999-10-21
NUMBER OF SEQ ID NOS: 30

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 700
TYPE: PRT
ORGANISM: Human
US-09-768-877-23

Query Match 46.2%; Score 48.5; DB 10; Length 700;
Best Local Similarity 52.6%; Pred. No. 1.2e+02;
Matches 10; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

OY 1 KKKLLKKPPPLKKLLK 19
DB 213 EWE-LKKPPNLFKIIK 230

RESULT 13
US-10-051-902-20

Sequence 20, Application US/10051902
Patent No. US20020178468A1
GENERAL INFORMATION:

APPLICANT: Allen, Steve
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tinney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/10/051,902
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 20
LENGTH: 513
TYPE: PRT
ORGANISM: Zea mays

Query Match 45.7%; Score 48; DB 9; Length 513;
Best Local Similarity 43.8%; Pred. No. 1e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WKLLKPPPLKKLL 17
DB 267 WKLLSPTPAVRRL 282

RESULT 14
US-10-051-909-20

Sequence 20, Application US/10051909
Patent No. US20020199217A1
GENERAL INFORMATION:

APPLICANT: Allen, Steve
APPLICANT: Helentjaris, Tim
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tinney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB1163 US CIP
CURRENT APPLICATION NUMBER: US/10/051,909
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083,044
PRIOR FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 20
LENGTH: 513
TYPE: PRT
ORGANISM: Zea mays

Query Match 45.7%; Score 48; DB 9; Length 513;
US-10-051-909-20

Best Local Similarity 43.8%; Pred. No. 1e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 WKLLKKPPPLKKLL 17
||||:|:|:|:
DB 267 WKLLSPTPAVRRL 282

RESULT 15

US-09-919-172-98
; Sequence 98, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 98
; LENGTH: 3256
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 2700132CD1
US-09-919-172-98

Query Match 45.7%; Score 48; DB 10; Length 3256;
Best Local Similarity 47.4%; Pred. No. 6.7e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 KWKKLLKKPPPLKKLLK 19
| | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 535 KRKSLVMTTPVWLKTIKE 553

Search completed: June 6, 2003, 13:26:03
Job time : 16.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 13:21:26 ; Search time 23.5 Seconds

(without alignments)
175.359 Million cell updates/sec

Title: US-10-081-418-2

Perfect score: 105

Sequence: 1 KMKRLKKPPPLKRLKRL 20

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriophage:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	50.5	357	5	O9GRQ3
2	52.5	50.0	173	10	O9FYP3
3	51	48.6	341	16	O8XMH7
4	51	48.6	369	5	O45769
5	50	47.6	900	10	O9LWM6
6	50	47.6	1437	5	O9GOS1
7	49	46.7	340	16	O9PNS4
8	49	46.7	656	3	O04373
9	48.5	46.2	608	5	O96849
10	48.5	46.2	608	5	O23898
11	48.5	46.2	700	4	O9HBB1
12	48.5	46.2	700	4	O8WU26
13	48.5	46.2	700	6	O9GIG1
14	48	45.7	552	13	O9YGC1
15	48	45.7	867	16	O9PKS5
16	47	44.8	110	15	O9WCA9

17	47	44.8	215	15	O9YIF1	O9YIF1 human immun
18	47	44.8	369	10	O9FMQ2	O9FMQ2 arabidopsis
19	47	44.8	524	15	O9ORL7	O9ORL7 porcine end
20	46.5	44.3	588	5	O9VFL0	O9VFL0 drosophila
21	46	43.8	85	17	O29172	O29172 archaeoglob
22	46	43.8	157	10	O39338	O39338 brassica na
23	46	43.8	373	16	O8R956	O8R956 thermocaneer
24	46	43.8	395	5	O9YD68	O9YD68 drosophila
25	46	43.8	529	10	O9FP52	O9FP52 oryza sativ
26	46	43.8	856	11	O9R2C5	O9R2C5 mus musculu
27	46	43.8	858	12	O9IBW9	O9IBW9 turkey herp
28	46	43.8	858	12	O9E6R1	O9E6R1 turkey herp
29	46	43.8	905	11	O9WUX5	O9WUX5 mus musculu
30	46	43.8	1341	3	O08281	O08281 saccharomyc
31	46	43.8	1342	3	O92271	O92271 saccharomyc
32	45.5	43.3	306	2	O93RKL	O93RKL haemophilus
33	45.5	43.3	355	17	O9V219	O9V219 pyrococcus
34	45.5	43.3	421	11	O8R486	O8R486 mus musculu
35	45.5	43.3	424	2	O93RK2	O93RK2 haemophilus
36	45.5	43.3	896	5	O94891	O94891 drosophila
37	45	42.9	120	15	O9O0Y5	O9O0Y5 porcine end
38	45	42.9	165	10	O29352	O29352 brassica na
39	45	42.9	299	4	O9Y785	O9Y785 homo sapien
40	45	42.9	325	13	O9PVL2	O9PVL2 gallus galli
41	45	42.9	342	10	O9FHJ5	O9FHJ5 arabidopsis
42	45	42.9	345	10	O9XJ2	O9XJ2 zea mays (m
43	45	42.9	345	12	O9DVS6	O9DVS6 pluteella xy
44	45	42.9	485	5	O95RE0	O95RE0 drosophila
45	45	42.9	493	10	O23213	O23213 arabidopsis

ALIGNMENTS

RESULT 1

ID O9GRQ3 PRELIMINARY; PRT; 357 AA.

AC O9GRQ3;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Hypothetical 39.6 kDa protein.

DE L6520.11.

GN Leishmania major.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5664;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIEDLIN;

RA Masny D., Purnelle B., Goffeau A., Ivens A.C., Quall M., Rajadream M.A., Barrell B.G.;

RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIEDLIN;

RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M., Smith D.F.;

RL "A physical map of the Leishmania major Friedlin genome.";

DR Genome Res. 8:135-145(1998).

DR EMBL; AL446005; CAC14538.1; -

KW Hypothetical protein.

SO SEQUENCE 357 AA; 39621 MW; D27C03EFP2B2139E CNC64;

Query Match 50.5%; Score 53; DB 5; Length 357;

Best local similarity 62.5%; Pred. No. 9.1;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 KMKPPPLKRLKRL 19

DB 85 KMKPPPLKRLKRL 100

```

RESULT 2
09FCP3 PRELIMINARY; PRT: 173 AA.
ID 09FCP3;
AC 09FCP3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE 50S ribosomal protein L29 (Putative 50S ribosomal protein L29).
AV565220.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Ericaceae; Ericaceae; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702.
RN
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones."
RL DNA Res. 5:203-216(1998).
RN
[2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene MON23.16/AT5G65220 (GI:10178184).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN
[3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
RA Toriumi M., Yu G., Brooks S., Chao O., Chen H., Karlin-Neumann G.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB013395; BAB1658.1;
DR EMBL: AY045966; AAK76640.1;
DR EMBL: AY079351; AAL85082.1;
DR InterPro: IPR001854; Ribosomal_L29.
DR Pfam: PF00831; Ribosomal_L29; 1.
DR TIGRfams: TIGR00012; L29; 1.
KM Ribosomal protein.
SQ SEQUENCE 173 AA; 19377 MW; B21C46FC9364CE42 CRC64;

Query Match 50.0%; Score 52.5; DB 10; Length 173;
Best Local Similarity 58.8%; Pred. No. 5.6;
Matches 10; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

OY 1 KMKK-LTKKPPPLTKKL 16
DB 141 QMKSTVPRPPSLKTL 157

```

```

OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL: AP003187; BAB80418.1;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 341 AA; 40334 MW; 196734ADCF0DC4E CRC64;

Query Match 48.6%; Score 51; DB 16; Length 341;
Best Local Similarity 58.8%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 KKLKKRPPLTKLTKL 19
DB 5 KKLKKRPVKKSKFKK 21

RESULT 4
045769 PRELIMINARY; PRT: 369 AA.
ID 045769;
AC 045769;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE T10C6.7 protein.
GN T10C6.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99065613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z93388; CAB07663.1;
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR PROSITE: PS50181; FBOX; 1.
SQ SEQUENCE 369 AA; 43506 MW; 00DEC66EAF291041 CRC64;

Query Match 48.6%; Score 51; DB 5; Length 369;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 KMKLTKKPPPLTKLTKL 20
DB 20 KFKRLQLPPPAFANILKIL 39

RESULT 5
09LYM6 PRELIMINARY; PRT: 900 AA.
ID 09LYM6;
AC 09LYM6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putable eukaryotic translation initiation factor 3 subunit 8
DE (Putable eukaryotic translation initiation factor 3 subunit 8).
GN F1802L110 OR AT3G56150.

```

```

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RA Benes V., Wurmbach E., Drzonek H., Ansoerge W., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Queller F., Salanoubat M.;
RA Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Yamada K., Banb J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shino P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA "Full length cDNA of gene At3g56150 (GI:1528840).";
RA Submitted (Nov-2001) to the EMBL/Genbank/DBJ databases.
RN
RA EMBL: AL163763; CAB87414.1; -.
DR EMBL: AY064006; AAL36362.1; -.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR000717; PCT.
DR Pfam: PF01399; PCT.1.
DR SMART: SM00088; PINT.1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Initiation factor.
SQ
SEQUENCE 900 AA; 102949 MW; 84EA72DDE49A05C5 CRC64;

Query Match 47.6%; Score 50; DB 10; Length 900;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 WKKLKKPPPLKLLKLLK 19
I:|:| | | | | | | |
Db 215 WEKMLSKDKLLKLLK 232

RESULT 6
O9G051 PRELIMINARY: PRT: 1437 AA.
AC O9G051;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Polypeptide.
GN POL.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN
RP SEQUENCE FROM N.A.
RA STRAIN-AX4; TRANSPOSON-GYPSY-LIKE LTR RETROTRANSPOSON DGLT-A1;
RA Glickner G., Saltzanski K., Winkler T., Dingermann T., Quail M.A.,
RA Cox E., Elchinger L., Noegel A.A., Rosenthal A.;
RA "The Complex Repeats of Dictyostelium discoideum.";
RA Genome Res. 11:585-594 (2001).
DR EMBL: AF298204; ANG37039.1; -.
DR InterPro: IPR001969; Aspartate_site.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; Rvse.
DR InterPro: IPR001878; ZnF_CCHC.
DR Pfam: PF00065; rve.1.
DR Pfam: PF00078; rvc.1.
DR Pfam: PF00098; zf-CCHC.1.
DR SMART: SM00343; ZnF_C2HC.1.

```

```

DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1437 AA; 164174 MW; 7242331E67AD9C6B CRC64;

Query Match 47.6%; Score 50; DB 5; Length 1437;
Best Local Similarity 55.0%; Pred. No. 83;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 WKKLKKPPPLKLLKLLK 20
I:|:| | | | | | | |
Db 666 KPNELKTPPSNCKELMKL 705

RESULT 7
O9PN54 PRELIMINARY: PRT: 340 AA.
AC O9PN54;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Uroporphyrinogen decarboxylase (EC 4.1.1.37).
DE HEME OR C11243.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN
RP SEQUENCE FROM N.A.
RA STRAIN-NCTC 11168;
RA MEDLINE-20150912; PubMed-1068204;
RA Partridge J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jørgensen K., Karslyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
RA reveals hypervariable sequences.";
RN
RT Nature 403:665-668 (2000).
DR EMBL: AL139077; CAB73497.1; -.
DR HSSP: P06132; URO.
DR InterPro: IPR000257; Uro_decarboxyls.
DR Pfam: PF01208; URO-D.1.
DR PROSITE: PS00906; UROD_1.1.
DR PROSITE: PS00907; UROD_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 340 AA; 38922 MW; 44137F84547C9PEC CRC64;

Query Match 46.7%; Score 49; DB 16; Length 340;
Best Local Similarity 55.0%; Pred. No. 32;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 WKKLKKPPPLKLLKLLK 20
I:|:| | | | | | | |
Db 160 KCKKMLYONPELKKLLKLLK 179

RESULT 8
O04373 PRELIMINARY: PRT: 656 AA.
AC O04373;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE D9719.2p.
GN YDR496C OR D9719.2.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,

```

RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
 RA Wiant A., Yelton M., Bolstein D., Davis R.W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Jia Y., Cherry J.M.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U33057; AAB64938.1; -.
 DR SGD: S0002904; YDR496C.
 DR InterPro: IPR001313; Pumlilo/Puf.
 DR Pfam: PF00806; PUF; 4.
 DR SMART: SM00025; Pumlilo; 6.
 SQ SEQUENCE 656 AA; 75105 MW; 4D172E8854326ECF CRC64;

Query Match 46.7%; Score 49; DB 3; Length 656;
 Best Local Similarity 47.4%; Pred. No. 58;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 WKRLKKPPPLKLLKLL 20
 Db 131 WERLRVKTPLPKQIREKL 149

RESULT 9

ID 096849 PRELIMINARY; PRT; 608 AA.
 AC 096849;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Slime mold (D.discoideum) transposon DIRS-1, complete, clone
 DE SB41.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX-3;
 RX MEDLINE=85061212; PubMed=6095047;
 RA Cappello J., Cohen S.M., Lodish H.F.;
 RT "Dictyostelium transposable element DIRS-1 preferentially inserts into
 RT DIRS-1 sequences.";
 RL Mol. Cell. Biol. 4:2207-2213(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX-3;
 RX MEDLINE=85085937; PubMed=6096693;
 RA Cohen S.M., Cappello J., Lodish H.F.;
 RT "Transcription of Dictyostelium discoideum transposable element DIRS-
 RT 1.";
 RL Mol. Cell. Biol. 4:2332-2340(1984).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX-3;
 RX MEDLINE=86079481; PubMed=2416457;
 RA Cappello J., Handelsman K., Lodish H.F.;
 RT "Sequence of Dictyostelium DIRS-1: An apparent retrotransposon with
 RT inverted terminal repeats and an internal circle junction sequence.";
 RL EMBL: M11339; AAA33195.1; -.
 DR InterPro: IPR000477; RVTse.
 DR Pfam: PF00078; RVT; 1.
 DR RNA-directed DNA polymerase.
 SQ SEQUENCE 608 AA; 70317 MW; BE89A7F0ED84AE73 CRC64;

Query Match 46.2%; Score 48.5; DB 5; Length 608;
 Best Local Similarity 47.6%; Pred. No. 63;

Matches. 10; Conservative 5; Mismatches 5; Indels 1; Gaps 1;
 Qy 1 KWKLKKPPP-LKKLLKLL 20
 Db 512 QWKQCLAFPPILLPILTEKM 532

RESULT 10

ID 023898 PRELIMINARY; PRT; 608 AA.
 AC 023898;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Hypothetical 70.4 kDa protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86079481; PubMed=2416457;
 RA Cappello J., Handelsman K., Lodish H.F.;
 RT "Sequence of Dictyostelium DIRS-1: an apparent retrotransposon with
 RT inverted terminal repeats and an internal circle junction sequence.";
 RL Cell 43:105-115(1985).
 DR EMBL: M11340; AAA70202.1; -.
 DR InterPro: IPR000477; RVTse.
 DR Pfam: PF00078; RVT; 1.
 KW Hypothetical protein; RNA-directed DNA polymerase.
 SQ SEQUENCE 608 AA; 70376 MW; B7285DAB3FE2FBF3 CRC64;

Query Match 46.2%; Score 48.5; DB 5; Length 608;
 Best Local Similarity 47.6%; Pred. No. 63;
 Matches 10; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 1 KWKLKKPPP-LKKLLKLL 20
 Db 512 QWKQCLAFPPILLPILTEKM 532

RESULT 11

ID 09HBBI PRELIMINARY; PRT; 700 AA.
 AC 09HBBI;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Calpain large polypeptide L2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN ASTROCYTOMA;
 RX MEDLINE=20403900; PubMed=10944468;
 RA Ye Z., Connor J.R.;
 RT "cDNA cloning by amplification of circularized first strand cDNAs
 RT reveals non-IRE-regulated iron-responsive mRNAs.";
 RL Biochem. Biophys. Res. Commun. 275:223-227(2000).
 DR EMBL: AF261089; AAF99682.1; -.
 DR HSSP: P04574; IALV.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001300; Protease_C2.
 DR InterPro: IPR000169; SHPTC_acetate.
 DR Pfam: PF01067; Calpain-III; 1.
 DR Pfam: PF00036; efhand; 3.
 DR Pfam: PF00648; Peptidase_C2; 1.
 DR PRINTS: PR00704; CALPAIN.
 DR SMART: SM00230; CysPC; 1.
 DR SMART: SM00054; EFh; 3.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.

RESULT 15

Q9PKS5

ID

PRELIMINARY;

PRT;

867 AA.

AC

Q9PKS5;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE ATP-dependent Clp protease, subunit B.

GN

TC0389.

OS Chlamydia muridarum.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_Taxid=83560;

RN

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=MOPN / NTG;

RX MEDLINE=20150255; Pubmed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,

RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,

RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,

RA Eisen J., Fraser C.M.;

RT "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia

RT pneumoniae AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

DR EMBL; AE002306; AAF39246.1;

DR TIGR; TC0389;

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003959; AAA_ATPase_cent.

DR InterPro; IPR001270; Chaperonin_clpA/B.

DR InterPro; IPR004176; Clp_N.

DR Pfam; PF00004; AAA; 1.

DR Pfam; PF02861; Clp_N; 2.

DR PRINTS; PR00300; CLPPROTEASEA.

DR SMART; SM00382; AAA; 1.

DR PROSITE; PS00870; CLPAB_1; 1.

KW Complete proteome.

SQ SEQUENCE 867 AA; 96603 MW; DAB4C579AAC79F57 CRC64;

Query Match

Best Local Similarity 45.7%; Score 48; DB 16; Length 867;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 WKLLKKPPPLKKLKL 20

DB 124 WKRTVKTSPKELIKL 142

Search completed: June 6, 2003, 13:24:54
 Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 13:21:00 ; Search time 7 seconds
(without alignments)
118.504 Million cell updates/sec

Title: US-10-081-418-2

Perfect score: 105

Sequence: 1 KWKLLKKPPPLKLLKLL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues 112892

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	47.6	900	IF38_ARATH	049160 arabidopsis
2	49	46.7	110	Y107_NPVAC	P41660 autographa
3	49	46.7	1235	KPB2_RABIT	P46018 oryctolagus
4	48.5	46.2	700	CAN2_HUMAN	P17655 homo sapien
5	48	45.7	292	V226_FOWPY	093509 fowlpox vir
6	48	45.7	339	DCUP_HELPJ	092108 helicobacte
7	48	45.7	339	DCUP_HELPJ	025325 helicobacte
8	47	44.8	1324	K167_HUMAN	P46013 homo sapien
9	46	43.8	366	POL2_GCMV	P13026 hungarian g
10	46	43.8	537	GAG_BAEVM	062618 drosophila
11	45.5	43.3	479	PBP4_HAEIN	P03341 baboon endo
12	45.5	43.3	700	CAN2_MOUSE	P45161 h penicilli
13	45.5	43.3	700	CAN2_MOUSE	008529 mus musculi
14	45.5	43.3	700	CAN2_MOUSE	007009 rattus norv
15	45	42.9	151	YPI6_METTM	P14305 methanobact
16	45	42.9	1148	YK00_CAEEL	P34305 caenorhabdi
17	44.5	42.4	629	STR_PYRHO	O59147 pyrococcus
18	44	41.9	143	NIFX_FRAL	O46038 frankia aln
19	44	41.9	354	DCUP_PASMU	P57964 pasteurella
20	44	41.9	591	PUR9_HUMAN	P31939 homo sapien
21	44	41.9	779	CC4_YEAST	P07834 saccharomyc
22	44	41.9	1310	ACR_RABIT	P12822 oryctolagus
23	43	41.0	244	Y293_MYGE	P47822 mycoplasma
24	43	41.0	487	MDM2_CANFA	P56950 canis famli
25	43	41.0	489	MDM2_MOUSE	P56950 mus musculi
26	43	41.0	491	MDM2_HORSE	P56951 equus cabal
27	43	41.0	491	MDM2_HUMAN	000987 homo sapien
28	43	41.0	613	UBID_VITBEH	094V48 vldrio choi
29	43	41.0	700	CAN2_CHICK	O92178 gallus gall
30	43	41.0	757	CC5_SCHPO	P39964 schizosacch
31	42.5	40.5	360	IME1_YEAST	P21190 saccharomyc
32	42	40.0	88	PLE4_BOVIN	P02777 bos taurus
33	42	40.0	152	UBC8_HUMAN	O14933 homo sapien

ALIGNMENTS

34	42	40.0	173	1	YNE5_CAEEL	P30644 caenorhabdi
35	42	40.0	282	1	TAND_ECOLI	P37510 escherichia
36	42	40.0	343	1	BIOB_BUCAL	P57378 buchnera ap
37	42	40.0	481	1	DDX6_XENLA	P54824 xenopus lae
38	42	40.0	483	1	DDX6_HUMAN	P26196 homo sapien
39	42	40.0	483	1	DDX6_MOUSE	P54823 mus musculi
40	42	40.0	494	1	TP6L_HUMAN	P08686 homo sapien
41	42	40.0	564	1	TP6L_PYRAB	O94135 pyrococcus
42	42	40.0	612	1	GIDA_MYGE	P47619 mycoplasma
43	42	40.0	1068	1	YCFO_MARPO	P12221 marchantia
44	42	40.0	1192	1	K2_DROME	O46072 drosophila
45	42	40.0	1199	1	Y173_HUMAN	O14679 homo sapien

RESULT 1
ID IF38_ARATH STANDARD: PRT; 900 AA.
AC 049160:

DT 30-MAY-2000 (rel. 39, Created)
DT 30-MAY-2000 (rel. 39, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Probable eukaryotic translation initiation factor 3 subunit 8 (eif3
p110) (p110).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99065153; PubMed=9849901;
RA Karniol B., Ianalom A., Kwok S., Tsuge T., Matsui M., Deng X.-W.,
RA Chamovitz D.A.;
RT "The Arabidopsis homologue of an eif3 complex subunit associates with
the COP9 complex.";
RT FEBS Lett. 439:173-179(1998).

CC - FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
SUBUNIT 1. EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS (BY
SIMILARITY).
CC - SIMILARITY: BELONGS TO THE EIF38 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sb.ch/announce/>
or send an email to license@isb-sb.ch)

CC EMBL: AF040102; AAC83464.1; -
CC InterPro: IPR000717; PCT.
CC Pfam: PF01399; PCT; 1.
CC SMART: SM00088; PINT; 1.
CC DR Initiation factor: Protein biosynthesis.
CC KW SEQUENCE 900 AA; 102989 MW; 3603FEDB59AB5D1 CRC64;
CC

Query Match 47.6%; Score 50; DB 1; Length 900;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 WKLLKKPPPLKLLKLL 19
DB 215 WEKLLSKDKLKLKLLK 232

RESULT 2
ID Y107_NPVAC STANDARD: PRT; 110 AA.
AC P41660;


```

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Met-|-Xaa or
CC Arg-|-Xaa with Leu or Val as the p2 residue.
CC -1- COFACTOR: Binds 3 calcium ions.
CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
CC of calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC membrane upon Ca++ binding.
CC -1- SIMILARITY: CONTAINS 5 EF-HAND CALCIUM-BINDING DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL; M23234; AAA35645.1; -
CC EMBL; J04700; AAA52760.1; -
CC PIR; A31218; A31218.
CC PDB; 1KRU; 07-DEC-01.
CC MEROPS; C02.002; -.
CC Genew; HGNC:1479; CAPN2.
CC MIM; 114230; -
CC DR InterPro; IPR002048; EF-hand.
CC DR InterPro; IPR001300; Protease_C2.
CC DR InterPro; IPR00169; SHprot_acsite.
CC DR Pfam; PF00036; efhand; 3.
CC DR Pfam; PF00648; peptidase_C2; 1.
CC DR Pfam; PF01067; Calpain_III; 1.
CC DR PRINTS; PR00704; CALPAIN.
CC DR ProDom; PD000012; EF-hand; 1.
CC DR SMART; SM00230; Cyspc; 1.
CC DR SMART; SM00054; Efh; 2.
CC DR PROSITE; PS00018; EF-HAND; 2.
CC DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
CC DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
CC DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
CC KW Hydroxylase: Thiol protease; Calcium-binding; Repeat; Multigene family;
CC 3D-structure.
CC FT PROPEP 1 19 ANCHORS TO THE SMALL SUBUNIT (POTENTIAL).
CC FT CHAIN 20 700 CALPAIN 2, LARGE (CATALYTIC) SUBUNIT.
CC FT DOMAIN 20 210 THIOL_PROTEASE DOMAIN I.
CC FT DOMAIN 211 355 THIOL_PROTEASE DOMAIN II.
CC FT DOMAIN 356 514 THIOL_PROTEASE DOMAIN III.
CC FT DOMAIN 515 529 LINKER.
CC FT DOMAIN 530 700 DOMAIN IV.
CC FT CA_BIND 541 552 EF-HAND 1.
CC FT CA_BIND 585 596 EF-HAND 2.
CC FT CA_BIND 615 626 EF-HAND 3.
CC FT CA_BIND 650 661 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
CC FT DOMAIN 680 691 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
CC FT ACT_SITE 105 105 BY SIMILARITY.
CC FT ACT_SITE 262 262 BY SIMILARITY.
CC FT ACT_SITE 286 286 BY SIMILARITY.
CC FT CONFLICT 68 68 S -> G (IN REF. 2).
CC FT CONFLICT 73 74 MR -> IE (IN REF. 2).
CC SO SEQUENCE 700 AA; 80006 MW; A944D13BC8465531 CRC64;

```

```

OY 1 KKKKKKKKKKKKKKK 19
Query Match 46.2%; Score 48.5; DB 1; Length 700;
Best Local Similarity 52.6%; Pred. No. 13;
Matches 10; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

```

```

DB 213 EWE-LKKPPNLEKTIKK 230

```

```

RESULT 5
V226_FOWPV

```

```

ID V226_FOWPV STANDARD; PRT; 292 AA.
AC 09J509;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine protein kinase FFW226 (EC 2.7.1.-).
GN FFW226.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tullman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus."
RL J. Virol. 74:3815-3831(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL; AF196100; AAF4570.1; -
CC InterPro; IPR000719; Euk_pkinase.
CC DR InterPro; IPR002290; Ser_thr_pkinase.
CC DR Pfam; PF00069; pkinase; 1.
CC DR ProDom; PD000001; Euk_pkinase; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC KW Transferase: Serine/threonine-protein kinase; ATP-binding.
CC FT DOMAIN 14 292 PROTEIN KINASE.
CC FT NP_BIND 20 28 ATP (BY SIMILARITY).
CC FT BINDING 43 43 ATP (BY SIMILARITY).
CC FT ACT_SITE 147 147 BY SIMILARITY.
CC SO SEQUENCE 292 AA; 33990 MW; CC0C0AB9632D3BD CRC64;

```

```

Query Match 45.7%; Score 48; DB 1; Length 292;
Best Local Similarity 47.1%; Pred. No. 6.5;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 4 KKKKKKKKKKKKK 20
DB 115 KKKKKKKKKKKKK 131

```

```

RESULT 6
DCUP_HELPJ STANDARD; PRT; 339 AA.
ID DCUP_HELPJ
AC 09ZLM8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
GN HEM6 OR JHP0551.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonghe B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;

```

```

RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- CATALYTIC ACTIVITY: Uroporphyrinogen-III - coproporphyrinogen + 4
CC CO(2).
CC -1- PATHWAY: Porphyrin biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001487; AAD06123.1; -.
DR HSSP: P06132; URO.
DR InterPro: IPR000257; Uro_decarbxy1s.
DR Pfam: PF01208; URO-D; 1.
DR PROSITE: PS00906; UROD_1; 1.
DR PROSITE: PS00907; UROD_2; 1.
DR Lyase; Decarboxylase; Porphyrin biosynthesis; Complete proteome.
SO SEQUENCE 339 AA; 38367 MW; 2BDDFF74D46C54 CRC64;

Query Match 45.7%; Score 48; DB 1; Length 339;
Best Local Similarity 55.0%; Pred. No. 7.6;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 KWKLLKKPPPLKLLKLL 20
   11111 11111111
Db 159 KSKKMLYSEPEVLKALLEKL 178

RESULT 7
DCUP_HELPY STANDARD; PRT; 339 AA.
ID DCUP_HELPY
AC 025325;
DT 15-JUN-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
GN HEME OR HP0604.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
CC Helicobacter.
CC NCBI_TaxID=210;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=26695 / ATCC 700392;
CC MEDLINE=97394467; PubMed=9252185;
CC Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
CC Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
CC Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
CC Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
CC McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
CC Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
CC Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
CC Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
CC Venter J.C.;
CC "The complete genome sequence of the gastric pathogen Helicobacter
CC pylori."
CC Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: Uroporphyrinogen-III - coproporphyrinogen + 4
CC CO(2).
CC -1- PATHWAY: Porphyrin biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000574; AAD07669.1; ALT_INIT.
DR HSSP: P06132; URO.
DR TIGR: HP0604; -.
DR InterPro: IPR000257; Uro_decarbxy1s.
DR Pfam: PF01208; URO-D; 1.
DR PROSITE: PS00906; UROD_1; 1.
DR PROSITE: PS00907; UROD_2; 1.
DR Lyase; Decarboxylase; Porphyrin biosynthesis; Complete proteome.
SO SEQUENCE 339 AA; 38339 MW; 76B03ADAA42726 CRC64;

Query Match 45.7%; Score 48; DB 1; Length 339;
Best Local Similarity 55.0%; Pred. No. 7.6;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 KWKLLKKPPPLKLLKLL 20
   11111 11111111
Db 159 KSKKMLYSEPEVLKALLEKL 178

RESULT 8
KI67_HUMAN STANDARD; PRT; 3256 AA.
ID KI67_HUMAN
AC P46013;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigen Ki-67.
GN MKI67.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC NCBI_TaxID=9606;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=94043435; PubMed=8227122;
CC Schlieter C., Duchrow M., Wohlenberg C., Becker M.H.G., Key G.,
CC Riad H.-D., Gerdes J.;
CC "The cell proliferation-associated antigen of antibody Ki-67: a very
CC large, ubiquitous nuclear protein with numerous repeated elements,
CC representing a new kind of cell cycle-maintaining proteins."
CC J. Cell Biol. 123:513-522(1993).
CC [2]
CC SEQUENCE OF 1-31 FROM N.A.
CC Gerdes J.;
CC Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THOUGHT TO BE REQUIRED FOR MAINTAINING CELL
CC PROLIFERATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. PREDOMINANTLY LOCALIZED IN THE G1
CC PHASE IN THE PERINUCLEAR REGION, IN THE LATER PHASES IT IS ALSO
CC DETECTED THROUGHOUT THE NUCLEAR INTERIOR, BEING PREDOMINANTLY
CC LOCALIZED IN THE NUCLEAR MATRIX. IN MITOSIS, IT IS PRESENT ON ALL
CC CHROMOSOMES.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION OF THIS ANTIGEN OCCURS
CC PREFERENTIALLY DURING LATE G1, S, G2, AND M PHASES OF THE CELL
CC CYCLE, WHILE IN CELLS IN G0 PHASE THE ANTIGEN CANNOT BE DETECTED.
CC -1- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X65550; CAA46519.1; -.

```

DR EMBL; X65551; CAA46520.1; -
 DR EMBL; X94762; CAA64388.1; -
 DR Genew; HGNC:7107; MK167.
 DR MIM:176741; -
 DR InterPro: IPR000253; FHA_domain.
 DR Pfam; PF00498; FHA; 1.
 DR SMART; SM00240; FHA; 1.
 DR PROSITE; PS50006; FHA_DOMAIN; 1.
 DR Cell cycle; Antigen; Nuclear protein; ATP-binding; Repeat;
 KW Alternative splicing.
 FT DOMAIN 27
 FT 1000 2928 FHA.
 FT REPEAT 1000 1112 16 X 122 AA APPROXIMATE REPEATS.
 FT REPEAT 1122 1234 1.
 FT REPEAT 1244 1356 2.
 FT REPEAT 1366 1477 3.
 FT REPEAT 1487 1598 4.
 FT REPEAT 1608 1720 5.
 FT REPEAT 1730 1842 6.
 FT REPEAT 1851 1964 7.
 FT REPEAT 1974 2086 8.
 FT REPEAT 2096 2204 9.
 FT REPEAT 2214 2326 10.
 FT REPEAT 2335 2447 11.
 FT REPEAT 2457 2569 12.
 FT REPEAT 2579 2689 13.
 FT REPEAT 2699 2808 14.
 FT REPEAT 2818 2928 15.
 FT NP_BIND 3034 3041 ATP (POTENTIAL).
 FT VASPLC 136 495 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 3256 AA; 358741 MW; 578F8C51BED42517 CRC64;

Query Match 45.7%; Score 48; DB 1; Length 3256;
 Best Local Similarity 47.4%; Pred. No. 74;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 KKKLKKPPPLKKLKK 19
 DB 535 KRKSLVMTPEVLKKIKE 553

RESULT 9
 POL2_GCMV STANDARD; PRT; 1324 AA.
 AC P13026;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE RN2 polypeptide (147 kDa protein) [Contains: Coat protein; 84 kDa protein].
 OS Hungarian grapevine chrome mosaic virus (GCMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
 OC Nepovirus.
 OC NCB1_TaxID=12273;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 814-822.
 RA MEDLINE-90016865; PubMed-2798129;
 RA Brault V., Hibrand L., Candresse T., Le Gall O., Dunez J.;
 RT "Nucleotide sequence and genetic organization of Hungarian grapevine chrome mosaic nepovirus RN2";
 RL Nucleic Acids Res. 17:7809-7823(1989).
 CC -1- FUNCTION: THE PROTEIN LOCATED AT THE N-TERMINAL END OF THE NEPOVIRUS POLYPROTEIN COULD BE REQUIRED FOR NEMATODE TRANSMISSION OF THE VIRUS.
 CC -1- PTM: THE RN2 POLYPROTEIN IS CLEAVED BY AN RN2-ENCODED PROTEASE TO YIELD THE MATURE COAT PROTEIN AND AN 84 KDA PROTEIN WHICH IS FURTHER CLEAVED INTO TWO PRODUCTS OF APPROXIMATELY 46 AND 48 KDA.
 CC -1- SIMILARITY: TO THE RN2 POLYPROTEIN OF OTHER NEPOVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X15163; CAA33255.1; -
 DR PIR; S06187; S06187.
 DR InterPro: IPR005054; Nepo_coat.
 DR InterPro: IPR005305; Nepo_coat_C.
 DR InterPro: IPR005306; Nepo_coat_N.
 DR Pfam; PF03688; Nepo_coat_C; 1.
 DR Pfam; PF03689; Nepo_coat_N; 1.
 DR Polypeptide; Coat protein.
 FT MOD_RES 811 811
 FT DOMAIN 811 810 BLOCKED (PROBABLE).
 FT 811 1324 84 KDA PROTEIN.
 FT VARIANT 400 400 N -> D.
 FT VARIANT 1006 1006 I -> T.
 SQ SEQUENCE 1324 AA; 147804 MW; 302C286B2EA1756A CRC64;

Query Match 44.8%; Score 47; DB 1; Length 1324;
 Best Local Similarity 52.9%; Pred. No. 42;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 WKLLKKPPPLKKLKK 18
 DB 1158 WFTLLRPDPKLSKILK 1174

RESULT 10
 M14A_DROME STANDARD; PRT; 366 AA.
 AC O62618; O62618; O9TYB4.
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase 14A (EC 2.7.1.-) (MAP kinase p38a) (p38 MAPK) (Dp38) (D-p38a).
 DE MPK3 OR P38A OR CG5A75.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, DEVELOPMENTAL STAGE, AND ENZYME REGULATION.
 RP TISSUE-Embryo;
 RC MEDLINE-98252940; PubMed-9584193;
 RA Han Z.S., Enslin H., Hu X., Meng X., Wu I.-H., Barrett T., Davis R.J., Ip Y.T.;
 RT "A conserved p38 mitogen-activated protein kinase pathway regulates Drosophila immunity gene expression";
 RL Mol. Cell. Biol. 18:3527-3539(1998).
 RN [2]
 RP SEQUENCE FROM N.A., PHOSPHORYLATION OF TYR-186, FUNCTION, SUBCELLULAR LOCATION, AND ENZYME REGULATION.
 RP TISSUE-Hemocyte;
 RC MEDLINE-98073070; PubMed-94117090;
 RA Han S.-J., Choi K.-Y., Brey P.T., Lee W.-J.;
 RT "Molecular cloning and characterization of a Drosophila p38 mitogen-activated protein kinase";
 RL J. Biol. Chem. 273:369-374(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RP STRAIN-Berkeley;
 RC MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA April J.F., Agbayan A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Betman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Boichan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Domnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelia S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Fiodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacible J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Teeter C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RA Science 287:2185-2195(2000).
 RL [4]
 RN SEQUENCE OF 38-188 FROM N.A.
 RP TISSUE-imaginal disks;
 RC MEDLINE=9235284; PubMed=1378625;
 RX Biggs W.H. III, Zipursky S.L.;
 RT "Primary structure, expression, and signal-dependent tyrosine
 phosphorylation of a *Drosophila* homolog of extracellular signal-
 regulated kinase.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:6295-6299(1992).
 RL CC
 CC -1- FUNCTION: Kinase involved in a signal transduction pathway. May
 down regulate insect immunity gene expression after prolonged
 infection.
 CC -1- ENZYME REGULATION: Activated by MKK3 and by phosphorylation on
 Tyr-186 in response to environmental stress.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
 CC Levels are highest at the preblastoderm stage but low levels are
 present throughout development.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF035546; AAC39030.1; -;
 DR EMBL; AF035547; AAC39031.1; -;
 DR EMBL; U86867; AAB97138.1; -;
 DR EMBL; AE003746; AAF56244.1; -;
 DR HSP; Q16539; IWC.
 DR FLYBase; FBgn0015765; MPK2.
 DR InterPro; IPR000719; Euk_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF000059; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00220; S_TKc; 1.

DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00113; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KW Nuclear protein; Phosphorylation.
 FT DOMAIN 25 312
 FT NP_BIND 31 39
 FT BINDING 54 54
 FT ACT_SITE 154 154
 FT MOD_RES 184 184
 FT MOD_RES 186 186
 FT CONFLICT 27 27
 FT CONFLICT 77 77
 FT CONFLICT 80 80
 FT CONFLICT 108 108
 FT CONFLICT 124 124
 FT CONFLICT 149 149
 FT CONFLICT 163 163
 SQ SEQUENCE 366 AA; 4225 MW; B3592B869F97990E CRC64;
 Query Match 43.8%; Score 46; DB 1; Length 366;
 Best Local Similarity 41.7%; Pred. No. 16;
 Matches 10; Conservative 4; Mismatches 4; Indels 6; Gaps 1;
 QY 1 KMKRLK-----KPPPLKRLK 18
 DB 340 KMKELIKETVTKRPPSYAQVLK 363
 ID GAG_BAEVM STANDARD; PRT; 537 AA.
 AC P03341; P10268;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GAG polypeptide [Contains: Inner coat protein P12; Core protein P15;
 DE Core shell protein P30; Nucleoprotein P10].
 GN GAG.
 OS Baboon endogenous virus (strain M7).
 OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_TaxID=11764;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kato S., Matsuo K., Nishimura N., Takahashi N., Takano T.;
 RT "The entire nucleotide sequence of baboon endogenous virus DNA: a
 RT chimeric genome structure of murine type C and simian type D
 RT retroviruses.";
 RL Jpn. J. Genet. 62:127-137(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83241915; PubMed=6408267;
 RA Tamura T.;
 RT "Provirus of M7 baboon endogenous virus: nucleotide sequence of the
 RT 969-pol region.";
 RL J. Virol. 47:137-145(1983).
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL
 POLYPEPTIDE.
 CC -1- SIMILARITY: CONTAINS 1 CCHC-TYPE ZINC FINGER.
 CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D10032; BAA00923.1; -;
 DR EMBL; X05470; CAA29027.1; -;

EMBL: J02034. -- NOT_ANNOTATED_CDS.
 DR PIR: A03939; FOMVVB.
 DR PIR: J0260; FOMVVB.
 DR InterPro: IPR000840; Gag_MA.
 DR InterPro: IPR003036; Gag_P30.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF00098; Zf-CCHC; 1.
 DR Pfam: PF01140; Gag_MA; 1.
 DR Pfam: PF02093; Gag_P30; 1.
 DR SMART: SM00343; Znf_C2HC; 1.
 DR PROSITE: PS50158; Zf_CCHC; 1.
 DR Core protein: Coat protein; Nucleoprotein; Polyprotein; Myristate;
 KM Zinc-finger.
 FT CHAIN 2 112 INNER COAT PROTEIN P12.
 FT CHAIN 113 226 CORE PROTEIN P15.
 FT CHAIN 227 477 CORE SHELL PROTEIN P30.
 FT CHAIN 478 537 NUCLEOPROTEIN P10.
 FT ZN_FING 501 518 CCHC-TYPE.
 FT LIPID 2 2 MYRISTATE.
 SQ SEQUENCE 537 AA; 60623 MW; 5DEE4437CFCEB79 CRC64;
 Query Match 43.8%; Score 46; DB 1; Length 537;
 Best Local Similarity 43.8%; Pred. No. 23;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 OY 2 WKLLKKPPLKKLL 17
 DB 86 MEDLVNPPVWVAPFL 101
 RESULT 12
 PBP4_HAEIN STANDARD; PRT; 479 AA.
 AC P45161;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Penicillin-binding protein 4 precursor (PBP-4) [Includes: D-alanyl-D-
 alanine carboxypeptidase (EC 3.4.16.4) (DD-peptidase) (DD-
 carboxypeptidase); D-alanyl-D-alanine-endopeptidase (EC 3.4.99.-) (DD-
 endopeptidase)].
 GN DNBC OR H11330.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
 OC Haemophilus.
 OX NCBI_Taxid=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Rd / KM20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kinkness E.F.,
 RA Kesteven A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McEwen K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Usterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd";
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: NOT INVOLVED IN TRANSEPTIDATION BUT EXCLUSIVELY
 CC CATALYZES A DD-CARBOXYPEPTIDASE AND DD-ENDOPEPTIDASE REACTION
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: D-alanyl-D-alanine + H(2)O -> 2 D-alanine.
 CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S13; ALSO KNOWN AS THE
 CC D-ALANYL-D-ALANINE CARBOXYPEPTIDASE 3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-stb.ch/announce/>
 CC or send an email to license@isb-stb.ch).
 CC -----
 DR EMBL: U32812; AAC22975.1; --
 DR MEROPS; S13.001; --
 DR TIGR: H11330; --
 DR InterPro: IPR000667; Peptidase_S13.
 DR Pfam: PF02113; Peptidase_S13; 1.
 DR PRINTS: PR00922; DADACPTASE3.
 DR TIGRFRMS: TIGR00666; PBP4; 1.
 KM Peptidoglycan synthetase; Cell division; Cell wall; Hydrolase; Signal;
 KW Antibiotic resistance; Periplasmic; Complete proteome.
 FT SIGNAL 1 26
 FT CHAIN 27 479 PENICILLIN-BINDING PROTEIN 4
 FT ACT_SITE 69 69 ACTIVATED BY PENICILLIN (FORMS PART OF
 FT ACT_SITE 420 420 THE ACTIVE-SITE CLEFT) (BY SIMILARITY).
 SQ SEQUENCE 479 AA; 52685 MW; 632868C6120CB48 CRC64;
 Query Match 43.3%; Score 45.5; DB 1; Length 479;
 Best Local Similarity 64.7%; Pred. No. 25;
 Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
 OY 4 KLLKP-PPLKKLLK 19
 DB 293 KLLSKPLPLKKMKMK 309
 RESULT 13
 CAN2_MOUSE STANDARD; PRT; 700 AA.
 AC O08529; O35518; O54843;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calpain 2, large [catalytic] subunit precursor (EC 3.4.22.17)
 DE (Calcium-activated neutral proteinase) (CANP) (M-type) (M-calpain)
 DE (Millimolar-calpain) (80 kDa M-calpain subunit) (CALP80).
 GN CAPR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c;
 RX MEDLINE=97480729; PubMed=9339374;
 RA Dear N., Matena K., Vinton M., Boehm T.;
 RA "A new subfamily of vertebrate calpains lacking a calmodulin-like
 RA domain: implications for calpain regulation and evolution";
 RL Genomics 45:175-184(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c;
 RA Ozaki Y.;
 RA submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CNS;
 RA Glass J.D., Nash N.R., Dry I., Culver D., Weseltingh S.;
 RA "Cloning of m-calpain from mouse nervous system";
 RA submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodelling and signal transduction (by similarity).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Met-I-Xaa or
 CC Arg-I-Xaa with Leu or Val as the p2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
 CC of calcium and inhibited by calpastatin.

```

CC      -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC      (regulatory) subunit.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC      membrane upon Ca++ binding.
CC      -1- SIMILARITY: CONTAINS 5 EF-HAND CALCIUM-BINDING DOMAINS.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: Y10139; CAAT1227.1; -.
DR      EMBL: D38117; BAA22964.1; -.
DR      EMBL: AF015038; AAB94029.1; -.
DR      HSSP: Q07009; IDP0.
DR      MEROPS: C02.002; -.
DR      MGD: MGI.88264; Capn2.
DR      InterPro: IPR002048; EF-hand.
DR      InterPro: IPR001300; Protease_C2.
DR      InterPro: IPR000169; SHprot_acstle.
DR      Pfam: PF000036; ehand. 3.
DR      Pfam: PF00648; peptidase_C2; 1.
DR      Pfam: PF01067; Calpain_III; 1.
DR      PRINTS: PR00704; CALPAIN.
DR      ProDom: PD000012; EF-hand; 1.
DR      SMART: SM00230; Cyspc; 1.
DR      SMART: SM00054; Efn; 2.
DR      PROSITE: PS00018; EF_HAND. 2.
DR      PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
DR      PROSITE: PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.
DR      PROSITE: PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
KW      Hydrolyase; Thiol protease; Calcium-binding; Repeat; Multigene family.
KW      PROPEP 1 19 ANCHORS TO THE SMALL SUBUNIT (POTENTIAL).
FT      CHAIN 1 20 700 CALPAIN 2, LARGE [CATALYTIC] SUBUNIT.
FT      DOMAIN 20 210 THIOL PROTEASE DOMAIN I.
FT      DOMAIN 211 355 THIOL PROTEASE DOMAIN II.
FT      DOMAIN 356 514 THIOL PROTEASE DOMAIN III.
FT      DOMAIN 515 529 DOMAIN III, C2-LIKE DOMAIN.
FT      DOMAIN 530 700 LINDER.
FT      CA_BIND 541 552 DOMAIN IV.
FT      CA_BIND 552 585 EF-HAND 1.
FT      CA_BIND 585 596 EF-HAND 2.
FT      CA_BIND 615 626 EF-HAND 3.
FT      DOMAIN 650 661 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
FT      DOMAIN 661 691 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT      ACT_SITE 680 691 BY SIMILARITY.
FT      ACT_SITE 105 105 BY SIMILARITY.
FT      ACT_SITE 262 262 BY SIMILARITY.
FT      ACT_SITE 286 286 BY SIMILARITY.
FT      CONFLICT 194 194 A -> T (IN REF. 1).
FT      CONFLICT 212 212 A -> G (IN REF. 2).
FT      CONFLICT 402 402 E -> G (IN REF. 1).
SQ      SEQUENCE 700 AA; 79871 MW; 682146B29068316 CRC64;
Query Match 43.3%; Score 45.5; DB 1; Length 700;
Best Local Similarity 47.4%; Pred. No. 36;
Matches 9; Conservative 6; Mismatches 3; Indels 1; Gaps 1;
QY      1 KWKLLKRPPLKLLKLLK 19
DB      213 EWEY-LRKPPPLFKIKK 230

```

```

DE      (Calcium-activated neutral proteinase) (CANP) (M-type) (M-calpain)
DE      (Molluscor-calpain).
GN      CAPN2.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_Taxid=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-94032492; PubMed-8218419;
RA      Deluca C.I., Davies P.L., Samis J.A., Elce J.S.;
RT      "Molecular cloning and bacterial expression of cDNA for rat calpain
RT      II 80 kDa subunit."
RL      Biochim. Biophys. Acta 1216:81-93(1993).
RN      [2]
RP      PARTIAL SEQUENCE.
RX      MEDLINE-21240297; PubMed-11342050;
RA      Moldoveanu T., Hosfield C.M., Jia Z., Elce J.S., Davies P.L.;
RT      "Ca(2+)-induced structural changes in rat m-calpain revealed by
RT      partial proteolysis."
RL      Biochim. Biophys. Acta 1545:245-254(2001).
RN      [3]
RP      MUTAGENESIS OF LYS-230; LYS-234 AND GLU-504.
RX      MEDLINE-21269273; PubMed-11102442;
RA      Hosfield C.M., Moldoveanu T., Davies P.L., Elce J.S., Jia Z.;
RT      "Calpain mutants with increased Ca2+ sensitivity and implications for
RT      the role of the C(2)-like domain."
RL      J. Biol. Chem. 276:7404-7407(2001).
RN      [4]
RP      MUTAGENESIS OF CYS-105; HIS-262; ASN-286 AND TRP-288.
RX      MEDLINE-95361909; PubMed-7635186;
RA      Arthur J.S., Gauthier S., Elce J.S.;
RT      "Active site residues in m-calpain: identification by site-directed
RT      mutagenesis."
RL      FEBS Lett. 368:397-400(1995).
RN      [5]
RP      X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX      MEDLINE-20069318; PubMed-10601010;
RA      Hosfield C.M., Elce J.S., Davies P.L., Jia Z.;
RT      "Crystal structure of calpain reveals the structural basis for
RT      Ca(2+)-dependent protease activity and a novel mode of enzyme
RT      activation."
RL      EMBO J. 18:6880-6889(1999).
CC      -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC      catalyze limited proteolysis of substrates involved in
CC      cytoskeletal remodelling and signal transduction.
CC      -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-xaa, Met-I-xaa or
CC      Arg-I-xaa with Leu or Val as the p2 residue.
CC      -1- COFACTOR: Binds 3 calcium ions.
CC      -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
CC      of calcium and inhibited by calpastatin.
CC      -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC      (regulatory) subunit.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC      membrane upon Ca++ binding.
CC      -1- SIMILARITY: CONTAINS 5 EF-HAND CALCIUM-BINDING DOMAINS.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L09120; AAI6327.1; -.
DR      PIR: S38361; S38361.
DR      PDB: IDP0; 2I-JUN-00.
DR      MEROPS: C02.002; -.
DR      InterPro: IPR002048; EF-hand.
DR      InterPro: IPR001300; Protease_C2.
DR      InterPro: IPR000169; SHprot_acstle.

```